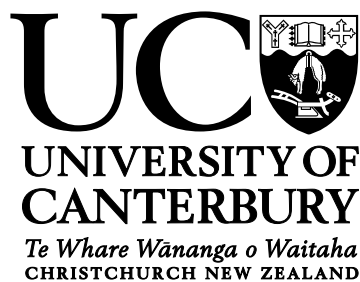


# **The Galactic Millisecond Pulsar Population**

## **Implications for the Galactic Center Excess**



**Harrison Ploeg**

School of Physical and Chemical Sciences  
University of Canterbury

This dissertation is submitted for the degree of  
*Doctor of Philosophy*

September 2021



## **Acknowledgements**

I would like to thank my supervisor Chris Gordon for his support and guidance throughout these last few years. His feedback and advice was always useful. I'd also like to thank Roland Crocker and Oscar Macias for their comments and contributions to the work that makes up this thesis.





## Abstract

Analysis of Fermi Large Area Telescope (Fermi-LAT) data has uncovered an extended gamma-ray source in the central region of the Milky Way. This Galactic Center Excess (GCE) has a spectral peak at a few GeV and appeared, initially, to have a spherically symmetric profile. These properties suggested that it may be evidence of self-annihilating weakly interacting massive particles (WIMPs) with a Navarro-Frenk-White (NFW) profile. However, the GCE gamma-ray spectrum is also similar to those of millisecond pulsars (MSPs) resolved in the Fermi-LAT data. An alternative possibility would then be that the GCE is produced by a population of MSPs in the Galactic Center too faint to be resolved individually as point sources at the present time. Providing further support for the MSP scenario, in recent years it has become apparent that the GCE may not be spherically symmetric, but may be spatially correlated with the distribution of stellar mass in the Galactic bulge.

In this thesis, we perform detailed modelling of the Galactic MSP population using data from the Fermi Large Area Telescope fourth source catalog data release 2 (4FGL-DR2) and the Australia Telescope National Facility (ATNF) pulsar catalog. Including in our model the spin down between formation and observation, we allow MSP luminosities to depend on intrinsic properties such as period  $P$ , magnetic field strength  $B$  and spectral energy cutoff  $E_{\text{cut}}$ . We find a model in which luminosity  $L \propto E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$  provides the best fit to the data, where  $a_\gamma = 1.2 \pm 0.3$ ,  $b_\gamma = 0.1 \pm 0.4$  and  $d_\gamma = 0.5 \pm 0.1$ , and where  $\dot{E} \propto B^2/P^4$  is the spin-down power. This model is significantly better than one in which luminosity is independent of other properties of MSPs, with the luminosity distribution being fitted directly. The Milky Way disk is expected to be the source of the resolved MSPs, with the GCE potentially produced by MSPs in the Galactic bulge. The Galactic bulge is divided into two structures: the boxy bulge, a significant bar structure extending a few kpc from the Galactic Center; and the nuclear bulge, a less massive component with radius  $\sim 230$  pc. Due to differing star formation histories it is expected that the MSPs in the Galactic bulge are older and therefore dimmer than those in the Galactic disk. Additionally, correlations between the spectral parameters of the MSPs and the spin-down rate of the corresponding neutron stars have been observed. This implies that the bulge MSPs may be spectrally different from the disk MSPs. Although we confirm

these correlations, we do not find they are sufficiently large to significantly differentiate the spectra of the bulge MSPs and disk MSPs when the uncertainties are accounted for. We find the age distributions of MSPs cannot be distinguished from a uniform birth rate, based on current data. Our results demonstrate that the population of MSPs that can explain the gamma-ray signal from the resolved MSPs in the Galactic disk and the unresolved MSPs in the boxy bulge and nuclear bulge can consistently be described as arising from a common evolutionary trajectory for some subset of astrophysical sources common to all these different environments. We do not require that there is anything systematically different about the inner Galaxy MSPs to explain the GCE. Additionally, we use a more accurate geometry for the distribution of bulge MSPs and incorporate dispersion measure estimates of the MSPs' distances. We find that the elongated boxy bulge morphology means that some bulge MSPs are closer to us and so easier to resolve. We identify three resolved MSPs that may belong to the bulge population.

In the “recycling” channel of MSP formation the neutron star forms from a core collapse supernovae that undergoes a random “kick” due to the asymmetry of the explosion. This would imply a smoothing out of the spatial distribution of the MSPs. We use  $N$ -body simulations to model how the MSP spatial distribution changes. We estimate the probability distribution of natal kick velocities using the resolved gamma-ray MSP proper motions, where MSPs have velocities relative to circular motion of  $\sigma_v = 77 \pm 6$  km/s. The scale of these peculiar velocities are determined as part of our Galactic MSP population model. We find that, due to the natal kicks, there is an approximately 10% increase in each of the bulge MSP spatial distribution dimensions and also the bulge MSP distribution becomes less boxy. We estimate that natal kicks change the axis ratios of the MSP distribution in the Galactic boxy bulge from  $\sim 1 : 0.43 : 0.40$  to  $\sim 1 : 0.46 : 0.44$ . Therefore, the bulge MSP distribution is still far from spherical.

# Table of contents

<b>List of figures</b>	<b>ix</b>
<b>List of tables</b>	<b>xi</b>
<b>Nomenclature</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The Galactic Bulge . . . . .	2
1.2 The Galactic Center Excess . . . . .	4
1.3 Millisecond Pulsars . . . . .	7
<b>2 Modelling the Galactic Millisecond Pulsar Population</b>	<b>13</b>
2.1 Introduction . . . . .	13
2.2 Method . . . . .	14
2.2.1 Modeling the Galactic Millisecond Pulsar population . . . . .	14
2.2.2 Fitting the model to data . . . . .	22
2.3 Results . . . . .	25
2.4 Discussion . . . . .	32
2.5 Conclusion . . . . .	55
<b>3 The Effect of Birth Kicks on the Distribution of Millisecond Pulsars</b>	<b>57</b>
3.1 Introduction . . . . .	57
3.2 Method . . . . .	58
3.3 Results . . . . .	64
3.4 Discussion and Conclusions . . . . .	66
<b>4 Summary and Future Work</b>	<b>89</b>
<b>References</b>	<b>93</b>

<b>Appendix A</b>	<b>Likelihood probability density function of resolved MSPs</b>	<b>105</b>
<b>Appendix B</b>	<b>Measurement Uncertainties</b>	<b>109</b>
<b>Appendix C</b>	<b>Sampling methods</b>	<b>111</b>
<b>Appendix D</b>	<b>Watanabe-Akaike Information Criterion (WAIC)</b>	<b>113</b>
<b>Appendix E</b>	<b>MWb and MWc0.8 <i>N</i>-body Simulation Figures</b>	<b>117</b>

# List of figures

1.1	Bar model of Freudenreich [1]. . . . .	3
1.2	Periods and period derivatives of pulsars. . . . .	8
2.1	Star formation rates for the disk, boxy bulge, and nuclear bulge. . . . .	17
2.2	Corner plot showing a selection of parameters for Model A1 ( $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ ) with 68% and 95% contours. . . . .	33
2.3	Corner plot showing a different selection of parameters for Model A1 with 68% and 95% contours. . . . .	34
2.4	Model A1 probability of delay time within each DTD bin. . . . .	35
2.5	Observed data compared to simulated observations for Model A1. . . . .	36
2.6	MSP luminosity distributions for Model A1. . . . .	37
2.7	MSP age distributions for Model A1. . . . .	38
2.8	Total observed gamma ray emission compared to all simulated MSP emission from inner $40^\circ \times 40^\circ$ region for Model A1. . . . .	39
2.9	Distribution of the number of disk and bulge MSPs for Model A1. . . . .	39
2.10	Distribution of the number of disk and bulge MSPs with $L > 10^{32} \text{ erg s}^{-1}$ for Model A1. . . . .	40
2.11	Distribution of the number of disk, boxy bulge and nuclear bulge MSPs produced per solar mass for Model A1. . . . .	40
2.12	Probability of observing $N$ bulge MSPs with the current sensitivity as well as for two and four times the current sensitivity for Model A1. . . . .	41
2.13	Comparison of posterior predictive boxy bulge GCE spectra for the $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ model in the case where we have a DTD (Model A1) and the case where a uniform MSP age distribution was used (Model A2). . . . .	42
2.14	Model A1 ( $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ ) posterior predictive corner plots. . . . .	43
2.15	Model A9 ( $L = \eta$ ) posterior predictive corner plots. . . . .	44
2.16	Model A7 ( $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , $a_{E_{\text{cut}}} = a_\Gamma = 0$ ) posterior predictive corner plots. . . . .	45

2.17	Comparison of correlation coefficients between $\log_{10}(P)$ and $\log_{10}(\text{DM})$ for resolved MSP simulated and real data for Model A1 ( $L = \eta E_{\text{cut}}^{a\gamma} B^{b\gamma} \dot{E}^{d\gamma}$ ) and Model A9 ( $L = \eta$ ). . . . .	46
2.18	Evolution of the luminosity and period distributions with time for Model A1. . . . .	47
2.19	Evolution of period with time for various magnetic field strengths $B$ , magnetic field axis angles $\alpha$ and initial periods $P_I$ . . . . .	48
2.20	Source count distribution for Model A1 MSP population. . . . .	50
2.21	Simulated spatial distribution of disk and bulge MSPs for Model A1. . . . .	55
3.1	Simulated galaxy rotation curves. . . . .	70
3.2	MW a model density maps. . . . .	71
3.3	MW b model density maps. . . . .	72
3.4	MW c0.8 model density maps. . . . .	73
3.5	MW a $x$ , $y$ and $z$ profiles with kicks at beginning. . . . .	74
3.6	MW a $x$ , $y$ and $z$ profiles with a uniform kick rate. . . . .	75
3.7	MW a flux distribution. . . . .	76
3.8	Simulated bar parameters with least square fit. . . . .	76
3.9	Line of sight contours for model of Galactic Center with predicted kicked version. . . . .	77
3.10	MW a $x$ , $y$ and $z$ 400 pc and 800 pc smoothed fitted bulge profiles with kicks at beginning. . . . .	81
3.11	MW a $x$ , $y$ and $z$ 400 pc and 800 pc smoothed fitted bulge profiles with a uniform kick rate. . . . .	82
3.12	MW a 400 pc and 800 pc smoothed bulge flux distribution . . . . .	83
3.13	MW a $x$ , $y$ and $z$ 200 pc smoothed fitted bulge profiles with kicks at beginning. . . . .	84
3.14	MW a $x$ , $y$ and $z$ 200 pc smoothed fitted bulge profiles with a uniform kick rate. . . . .	85
3.15	MW a 200 pc smoothed bulge flux distribution . . . . .	86
3.16	MW a $x$ , $y$ and $z$ profiles with kicks at beginning for various kick scales. . . . .	87
3.17	MW a $x$ , $y$ and $z$ profiles with a uniform kick rate for various kick scales. . . . .	88
E.1	MW b $x$ , $y$ and $z$ profiles with kicks at beginning. . . . .	118
E.2	MW b $x$ , $y$ and $z$ profiles with a uniform kick rate. . . . .	119
E.3	MW c0.8 $x$ , $y$ and $z$ profiles with kicks at beginning. . . . .	120
E.4	MW c0.8 $x$ , $y$ and $z$ profiles with a uniform kick rate. . . . .	121
E.5	MW b flux distribution. . . . .	122
E.6	MW c0.8 flux distribution. . . . .	122

# List of tables

2.1	Prior ranges for MSP population model parameters. . . . .	28
2.2	Average WAIC for each model. . . . .	29
2.3	Medians and 68% intervals for model parameters of a selection of MSP population models. . . . .	31
2.4	Correlation coefficient posterior predictive p-values between observables for model A1. . . . .	37
2.5	Correlation coefficient posterior predictive p-values between observables for model A6. . . . .	38
2.6	Correlation coefficient posterior predictive p-values between observables for model A7. . . . .	41
2.7	Correlation coefficient posterior predictive p-values between observables for model A9. . . . .	42
2.8	Details of MSPs with probability greater than 5% of coming from bulge population for Model A1. . . . .	54
3.1	Kick velocity Maxwell distribution parameters that produce a peculiar velocity distribution consistent with resolved gamma-ray MSPs. . . . .	66
3.2	Model MWa fitted parameters. . . . .	67
3.3	Model MWb fitted parameters. . . . .	68
3.4	Model MWc0.8 fitted parameters. . . . .	69
3.5	Change in $-2\log(L)$ using different bar models. . . . .	69
3.6	Least square fit values for simulated bar parameters. . . . .	70
3.7	Prediction of Milky Way bar kicked spatial distribution. . . . .	70





# Nomenclature

## Acronyms / Abbreviations

4FGL-DR2 Fermi Large Area Telescope fourth source catalog data release 2

AIC Accretion Induced Collapse

ATNF Australia Telescope National Facility

COBE Cosmic Background Explorer

DIRBE Diffuse Infrared Background Experiment

DM Dispersion Measure

DTD Delay Time Distribution

Fermi-LAT Fermi Large Area Telescope

GCE Galactic Center Excess

LMXB Low Mass X-ray Binary

MCMC Markov Chain Monte Carlo

MSP Millisecond Pulsar

NFW Navarro-Frenk-White

NSC Nuclear Stellar Cluster

NSD Nuclear Stellar Disk

OGLE Optical Gravitational Lensing Experiment

SFR Star Formation Rate

VVV VISTA Variables in the Via Lactea

WAIC Watanabe Akaike Information Criterion

WIMP Weakly Interacting Massive Particle

# Chapter 1

## Introduction

The Galactic Center Excess (GCE) is an extended gamma-ray source detected in the Fermi Large Area Telescope (Fermi-LAT) data in the central region of the Galaxy. This source, which has a spectral peak at a few GeV, initially appeared to have a spherically symmetric density profile, suggesting that it may be evidence of dark matter self-annihilating in the form of weakly interacting massive particles (WIMPs) with a Navarro-Frenk-White (NFW) distribution [2–5]. More recently, however, more detailed examination has revealed that the GCE may not be spherically symmetric but exhibits a spatial morphology that is correlated with the distribution of stellar mass in the Galactic bulge [6–10]. One recent study, however, using different methods, still argues for a spherically symmetric GCE [11]. If the GCE does trace the stellar mass of the bulge this would disfavor a dark matter origin and would point to a scenario in which it is produced by a population of dim, unresolved, astrophysical point sources such as Millisecond Pulsars (MSPs) [12]. Millisecond pulsars are rapidly spinning neutron stars with millisecond periods which emit gamma radiation with a spectrum that also peaks at a few GeV. There is some debate about whether the resolved MSPs are consistent with the needed bulge population, see for example refs. [13–16].

In this thesis we focus on the Galactic population of MSPs and explore whether observations could be consistent with a GCE produced by an unresolved population of MSPs. In Chapter 2, based on Ploeg et al. [17], we model the Galactic population of MSPs using the resolved gamma-ray MSPs detected by Fermi-LAT. Extending the model of Ploeg et al. [15], where the only intrinsic property of pulsars was their luminosity, we now include other properties such as the period, magnetic field strength, and spectral parameters, upon which the luminosity may depend [18, 19]. In Chapter 2 the bulge MSPs that could produce the GCE are assumed to differ only in their star formation rate (SFR) and so we can, under the assumption that the GCE is entirely produced by MSPs, estimate the total number of

bulge MSPs we may expect to have already resolved. Fitting several models to data, we find the best fit is one in which the luminosities of MSPs depend on their spectral energy cutoff; their magnetic field strength; and their spin-down power, the rate at which the pulsar loses rotational kinetic energy. We find that Galactic bulge MSPs do not need to be systematically different to explain the GCE, which could be produced by an inner Galaxy population of a few tens of thousands. In Chapter 3, based on Ploeg and Gordon [20], we investigate the effect of birth kicks on the spatial distribution of MSPs, with a particular interest in the effect on the Galactic bulge distribution. We do this by running  $N$ -body simulations intended to approximate the Milky Way and introduce a distribution of particles which are given a random Maxwell distributed kick of an appropriate scale. We then fit a parametric model to the distributions of particles with no kick, kicks occurring at the beginning of the simulation, and kicks occurring at a uniform rate throughout the simulation. We find that the boxy bulge structure in the center of simulated galaxies is broadened slightly by kicks and becomes less boxy, with the scale parameters increasing by approximately 10%, however, it remains far from spherical. In the remainder of this chapter we provide further background on the Galactic bulge, GCE, and MSPs.

## 1.1 The Galactic Bulge

The inner region of the Milky Way has a significant bar or boxy bulge structure. This structure can be found in near infrared imagery from the Diffuse Infrared Background Experiment (DIRBE) of the Cosmic Background Explorer (COBE) after accounting for emission from material between the Sun and the Galactic Center, and also extinction associated with dust [1, 21–23]. For example, Freudenreich [1] fitted a 47 parameter model of the Milky Way to 1.25, 2.2, 3.5 and 4.9  $\mu\text{m}$  images. This model included a warped disk with a central hole, a dust density model of a similar form, and three alternative bar models. In addition to contributing to the modelled integrated emission along lines of sight, the dust model was used to calculate an extinction correction factor  $\exp(-\tau_v(s))$  at a distance  $s$ . The best bar model was of the form:

$$\rho_{\text{F98, bar}} \propto \text{sech}^2(R_s) \quad (1.1)$$

where:

$$R_{\perp}^{C_{\perp}} = \left(\frac{|x|}{a_x}\right)^{C_{\perp}} + \left(\frac{|y|}{a_y}\right)^{C_{\perp}} \quad (1.2)$$

$$R_s^{C_{\parallel}} = R_{\perp}^{C_{\parallel}} + \left(\frac{|z|}{a_z}\right)^{C_{\parallel}} \quad (1.3)$$

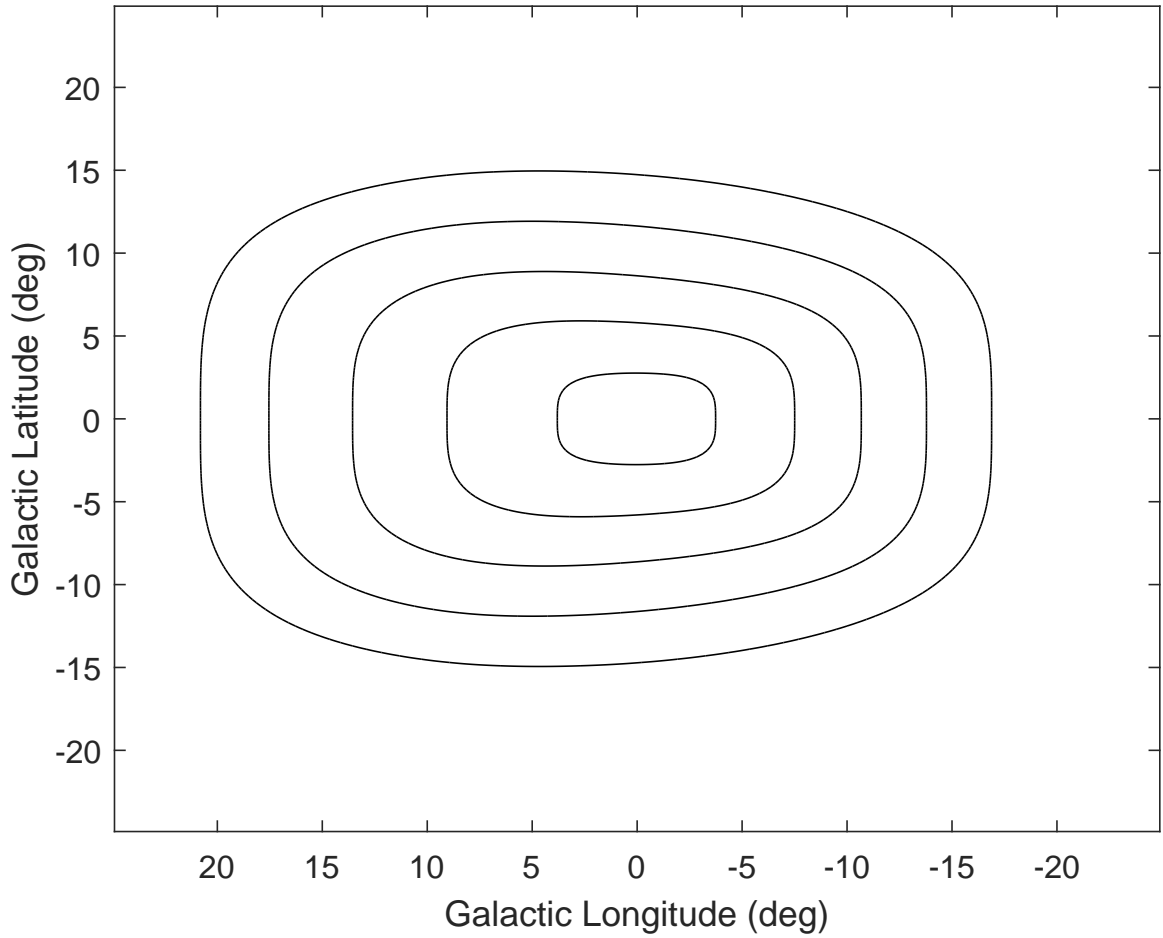


Fig. 1.1 Integrated density along lines of sight for Freudenreich [1] bar model. The contours are at  $1/16$ ,  $1/4$ ,  $1$ ,  $4$  and  $16$  times the mean integrated density in this region.

and where  $x$ ,  $y$  and  $z$  are coordinates in the bar frame, which may be rotated to determine the angle between the location of the Sun and the bar major axis along  $x$ . The parameters  $a_x$ ,  $a_y$  and  $a_z$  are the scale lengths along  $x$ ,  $y$  and  $z$  respectively. The face-on and edge-on shape parameters,  $C_{\perp}$  and  $C_{\parallel}$ , lead to a diamond shape when less than 2 and a boxy shape when greater. Assuming the distance to Galactic Center is 8.5 kpc, Freudenreich found  $a_x = 1.7$  kpc,  $a_y = 0.64$  kpc and  $a_z = 0.44$  kpc. With  $C_{\perp} = 1.6$  and  $C_{\parallel} = 3.5$ , the bar has a boxy shape when viewed edge-on. The angle of the bar relative to the Sun was  $13.8^{\circ}$ . The contour plot of Fig. 1.1 shows this bar as viewed from the location of the Sun, where we have integrated along lines of sight through the density model. The boxy appearance is clear. Using red clump giants, work using observations from the Optical Gravitational Lensing Experiment (OGLE) [24–26] and the VISTA Variables in the Via Lactea (VVV) survey [27, 28] have come to similar conclusions. As red clump giants have a relatively

narrow luminosity distribution, estimates of their distance can be made using their apparent magnitude. Cao et al. [26] use a boxy bulge model with density:

$$\rho_{\text{Cao13, bar}} \propto K_0(R_s) \quad (1.4)$$

where  $K_0$  is the modified Bessel function of the second kind and with fixed parameters  $C_\perp = 2$  and  $C_\parallel = 4$ . They find the bar is at an angle of 29.4 deg and has scale parameters of 0.67 kpc, 0.29 kpc and 0.27 kpc along  $x$ ,  $y$  and  $z$  respectively. Combining red clump giant data from multiple surveys, Wegg et al. [29] find a total bar half length of  $5.0 \pm 0.2$  kpc. As part of this structure, they also find evidence for a two component long bar, a thin extension to the boxy bulge. The scale heights of these two components are  $\approx 180$  pc and  $\approx 45$  pc, and they appear to be aligned with the boxy bulge with an angle of 28 to 33 degrees. Overall, these studies of the Galactic bulge tend to find the sun is located at an angle  $\lesssim 30$  deg from the major axis of the bar.

A second, much smaller, structure present at the Galactic Center is the nuclear bulge. This consists of two components: the Nuclear Stellar Cluster (NSC), and the Nuclear Stellar Disk (NSD). The NSC, with a mass of  $\sim 3 \times 10^7 M_\odot$ , has an  $r^{-2}$  density profile in the inner few parsecs, while the far more massive NSD,  $M_{\text{NSD}} = (1.4 \pm 0.6) \times 10^9 M_\odot$ , is a disk with radius  $\sim 230$  pc and scaleheight of  $45 \pm 5$  pc [30].

If the GCE is produced by a population of unresolved astrophysical sources, such as MSPs, we expect that those sources would have a spatial distribution similar to that of the Galactic bulge stellar mass. In Chapter 2 the Freudenreich [1] model of the boxy bulge (Eq. 1.1) is used to model the bar population of MSPs; a parametric model described in that chapter is used for the nuclear bulge population. In Chapter 3 a model of the same form as the Cao et al. [26] boxy bulge (Eq. 1.4) is used as a component of the parametric model fitted to the spatial distribution of particles at the end of  $N$ -body simulations.

## 1.2 The Galactic Center Excess

The Large Area Telescope (LAT) is a gamma-ray telescope on board the Fermi Gamma-ray Space Telescope (Fermi) [31]. Fermi-LAT, which images the entire sky approximately every 3 hours, detects gamma-rays when they convert to electron-positron pairs upon interacting with a thin layer of tungsten foil. The energy of the gamma-ray is estimated by measuring the energy of the charged particles produced and tracking them allows the incoming direction to be reconstructed. It is sensitive to gamma-rays in the range of 20 MeV to 300 GeV, with

uncertainty in direction declining as energy increases, decreasing from an angular resolution of a few degrees at 100 MeV to  $\sim 0.1$  deg at 100 GeV [32].

Based on about a year of gamma-ray data from Fermi, Goodenough and Hooper [2] reported that the radiation from the inner few degrees around the center of the Galaxy could be consistent with annihilating dark matter with an inner  $r^{-\gamma}$  profile with  $\gamma = 1.1$ . Further studies [3–5, 33–36] confirmed the presence of this apparently spherically symmetric extended source of gamma-rays with  $\gamma \approx 1.2$  and with a spectrum peaking at a few GeV. The spatial distribution of the dark matter halo can be modelled using:

$$\rho(r) = \frac{\rho_0}{\left(\frac{r}{r_s}\right)^\gamma \left(1 + \left(\frac{r}{r_s}\right)^\alpha\right)^{(\beta-\gamma)/\alpha}} \quad (1.5)$$

which is a generalized version of the Navarro-Frenk-White (NFW) profile where  $\alpha = 1$ ,  $\beta = 3$  and  $\gamma = 1$  [37]. The behavior near the Galactic Center where  $r \ll r_s$  is  $\rho \propto r^{-\gamma}$ . The dark matter annihilation flux as a function of energy  $E_\gamma$  and galactic coordinates  $l$  and  $b$  can be written as [5, 38]:

$$\Phi(E_\gamma, b, l) = \Phi^{\text{PP}}(E_\gamma) \times J(b, l) \quad (1.6)$$

where  $\Phi^{\text{PP}}$  is a particle physics contribution:

$$\Phi^{\text{PP}} = \frac{1}{2} \frac{\langle \sigma v \rangle}{4\pi m_{\text{DM}}^2} \sum_f \frac{dN_f}{dE_\gamma} B_f \quad (1.7)$$

where  $\langle \sigma v \rangle$  is the mean self-annihilation cross-section multiplied by the relative velocity of the dark matter particles,  $m_{\text{DM}}$  is the WIMP mass,  $dN_f/dE_\gamma$  is the spectrum produced by annihilation channel  $f$  with branching ratio  $B_f$ . The factor  $J(b, l)$  accounts for the dark matter profile along the line of sight at  $l$  and  $b$ , where the flux produced at a point in space is proportional to the square of the density of dark matter particles:

$$J(b, l) = \int ds \rho^2(r(s, b, l)) \quad (1.8)$$

Hooper and Goodenough [39] found 7–10 GeV WIMPs annihilating to the tau lepton pair  $\tau^+ \tau^-$  provided a good fit to the data. Abazajian and Kaplinghat [4] find annihilations of 10 GeV to 1 TeV WIMPs to  $b\bar{b}$  quarks and 10–30 GeV WIMPs to  $\tau^+ \tau^-$ . Gordon and Macias [5] suggest WIMPs with mass 20–60 GeV annihilating to a mixture of  $\tau^+ \tau^-$  and  $b\bar{b}$ . In addition to finding the spectrum and spatial morphology of the GCE is consistent with self-annihilating WIMPs distributed according to an NFW profile, these studies also find annihilation cross-sections similar to the thermal relic value of  $\langle \sigma v \rangle \approx 2.2 \times 10^{-26} \text{ cm}^3 \text{ s}^{-1}$ ,

the predicted value, given the current dark matter density, for WIMPs of mass  $\gtrsim 10$  GeV frozen out of thermal equilibrium in the early universe as it cooled [40].

An alternative possibility is a population of point sources in the Galactic Center unresolved by Fermi-LAT, such as MSPs which have similar gamma-ray spectra to the GCE [12]. The possibility of a population of MSPs in the inner Galaxy was proposed as early as 2005 in order to explain diffuse gamma-ray emission observed by EGRET in the Galactic Center region [41]. As the flux would be proportional to their density, they would instead have a density distribution  $\rho(r) \propto r^{-2\gamma}$ . Brandt and Kocsis [42] proposed that this profile could have arisen from the disruption of globular clusters. During this process, dynamical friction causes the orbits of globular clusters to decay and tidal forces disrupt the cluster as it approaches the Galactic Center. According to Gnedin et al. [43] this would lead to a distribution of stars, including MSPs, with inner profile  $\rho(r) \sim r^{-2.2}$ .

Yuan and Zhang [44] used resolved gamma-ray MSP data from the second Fermi-LAT catalog of gamma-ray pulsars [45] to show that a population of MSPs with an  $r^{-2.4}$  profile and with a luminosity distribution similar to those in the disk could explain the GCE without any being resolved. A similar analysis was performed by Petrovic et al. [46], who also suggested the possibility that there may be a secondary emission contribution to the GCE caused by inverse Compton scattering by relativistic leptons escaping MSPs. Hooper and Mohlabeng [13] later argued, fitting the spatial and luminosity distribution of MSPs in the Fermi-LAT third source catalog [47], that between 15 and 43 Galactic Center MSPs would have been resolved. However, Ploeg et al. [15] found, using different methods, that there existed a region of parameter space that was consistent with the observed MSP data and where a population of bulge MSPs with the same luminosity distribution could explain the GCE while having a high probability that none were resolved. Hooper and Mohlabeng [13] performed a binned fit to resolved MSP Galactic coordinates and flux, whereas Ploeg et al. [15] used an unbinned fit and also included parallax distance estimates where available.

There is statistical evidence of unresolved point sources near the Galactic Center below the Fermi-LAT detection threshold that may explain at least a significant fraction of the GCE [48–51]. In addition, in recent years it has become increasingly apparent that the GCE may in fact be correlated with the distribution of stellar mass in the Galactic Center [6–10]. These studies perform fits using spatial templates modelling the contributions from sources distributed as the boxy bulge and nuclear bulge, as well as from NFW squared templates modelling dark matter annihilation. They do not find a significant NFW squared component, strongly disfavoring a scenario in which a spherically symmetric distribution of annihilating dark matter is the dominant source of the GCE. Bartels et al. [7] find the best fitting boxy bulge model of Cao et al. [26] (Eq. 1.4) provided a good fit to the data. Using



the Freudenreich [1] model (Eq. 1.1), Macias et al. [8] argue for an admixture formation scenario where there is a mixture of primordial and dynamical formation. In the primordial case, the MSPs would simply follow the bar density  $\rho_{\text{bar}}$ , whereas for dynamical formation their density would be proportional to the stellar encounter rate which in turn is proportional to the bar density squared,  $\rho_{\text{bar}}^2$ . Adding a parameter  $s$ , they find using the density  $\rho_{\text{F98, bar}}^s$ , a best fit value of  $s = 1.4$ . They also show, by fitting a combination of the primordial and dynamical templates, the primordial channel contribution is  $(52 \pm 23)\%$ . These results are dependent on modelled Galactic diffuse gamma-ray emission; for example, Di Mauro [11] recently argued, based on 11 years of Fermi-LAT data, that the spatial morphology of the GCE remains consistent with dark matter annihilating with a spherically symmetric NFW profile.

In summary, while WIMPs annihilating to gamma-rays cannot at present be ruled out, the evidence for sub-threshold point sources and a non-spherical spatial distribution suggests that the GCE may be produced by a population of unresolved astrophysical point sources. The spectrum of the GCE is also similar to Fermi-LAT detected MSPs, and it is these we consider in this thesis.

## 1.3 Millisecond Pulsars

Pulsars are rapidly rotating neutron stars. Their radio and gamma-ray emission is driven by the acceleration of charged particles in strong magnetic fields. This emission is not isotropic, with the strongest emission along the magnetic axis leading to pulsations as the pulsar rotates. They can be divided into two groups: MSPs, and young pulsars. MSPs are pulsars with rotation periods  $P$  on the order of milliseconds, they also tend to spin down at a slower rate, with  $\dot{P} \lesssim 10^{-18}$ . In Fig. 1.2, the periods and rate of spindown of pulsars in the Australia Telescope National Facility (ATNF) pulsar catalog are shown, the MSPs are clearly seen in the bottom left corner separated by a region of relatively few pulsars from the larger population of young pulsars. For MSPs,  $\dot{P}$  is small enough that a correction should be applied, where the true period rate of change is [36]:

$$\dot{P}_{\text{int}} = \dot{P} - \Delta\dot{P} \quad (1.9)$$

where  $\Delta\dot{P}$  accounts for the Shklovskii effect [53] and radial acceleration in the Galactic potential. This correction of order  $\sim 10^{-21}$  will be important later in Chapter 2. The surface magnetic field strength of a pulsar  $B \propto \sqrt{P\dot{P}}$  is much smaller for MSPs ( $\sim 10^8$ – $10^9$  G) than young pulsars ( $\sim 10^{11}$ – $10^{13}$  G). Assuming pulsars spin down via magnetic dipole braking

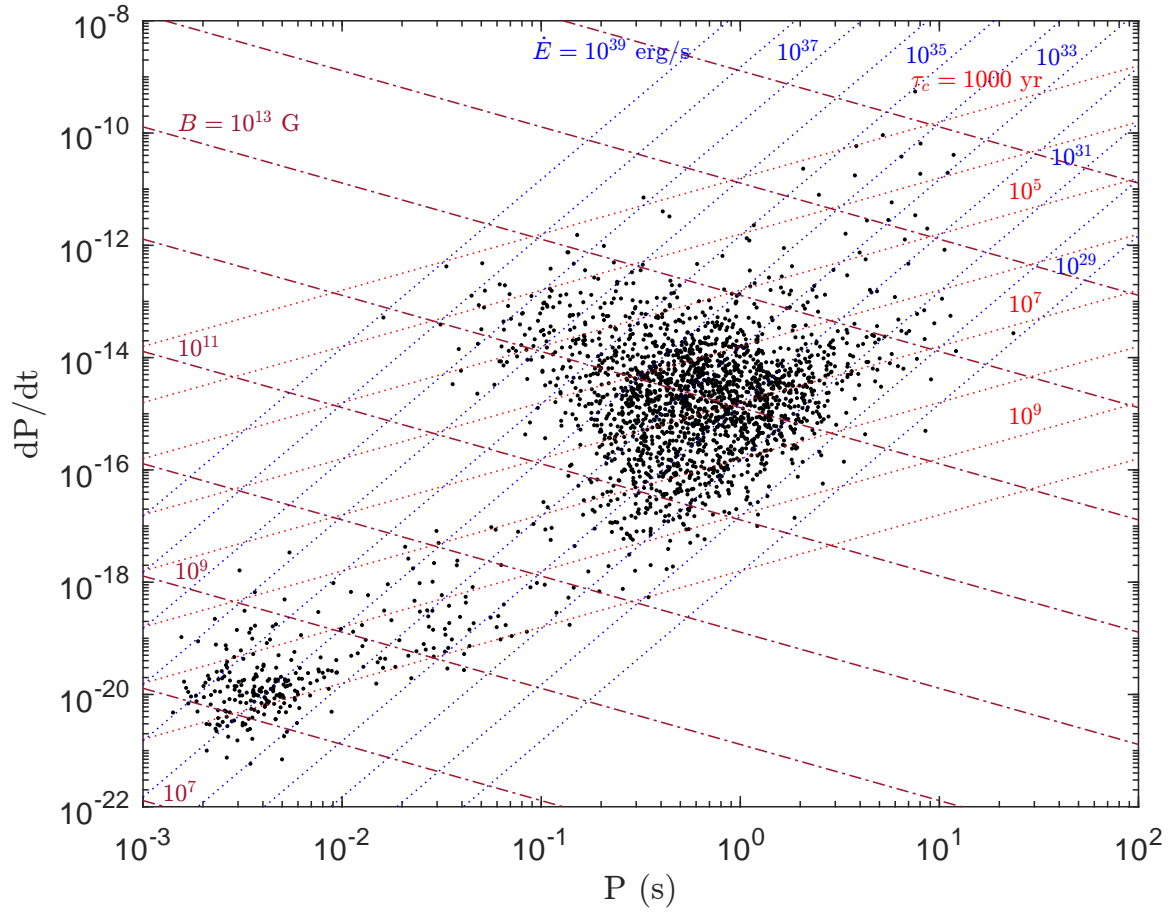


Fig. 1.2 Periods  $P$  and measured period derivatives  $\dot{P}$  of pulsars in ATNF catalog [52]. Included are lines of constant magnetic field strength  $B$ , spin-down power  $\dot{E}$  and characteristic age  $\tau_c$ .

and that the birth period  $P_l \ll P$ , we can define the characteristic age as [54]:

$$\tau_c = P/2\dot{P} \quad (1.10)$$

For MSPs these are  $\sim 1\text{--}10$  Gyr and for young pulsars  $\lesssim 100$  Myr. As a result of their far weaker magnetic fields, the MSPs spin down very slowly when compared to young pulsars. This is part of the reason we do not further consider young pulsars as a likely explanation for the GCE, a lack of recent star formation means they would have long ago spun down to the point where they no longer produce sufficient gamma radiation [15]. However, O’Leary et al. [55] have argued there may be enough star formation in the inner  $\lesssim 200$  pc for them to make a substantial contribution.

The “recycling” model of MSP formation involves the accretion of material from a binary companion onto an old pulsar formed in a core collapse supernova, transferring angular momentum and spinning up the pulsar to millisecond periods [56–58]. This process can occur when, assuming the binary system remains intact after the supernova, the companion evolves and overfills its Roche lobe, losing mass to the neutron star. Given the much stronger magnetic fields of young pulsars, this model of MSP formation also requires that there is some mechanism by which the magnetic field strengths decay to around the  $10^8\text{--}10^9$  G range. For example, if the magnetic field is generated by currents in the neutron star crust, one possibility is that decay of the magnetic field may be accelerated by heating of the crust during the accretion process. When the crust is heated, its conductivity is decreased and as a result the rate of Ohmic dissipation of currents is increased [59, 60]. During the accretion phase the system could potentially be seen as a low mass X-ray binary (LMXB) and Haggard et al. [14] have argued, assuming the ratio of total MSP luminosity in globular clusters to the number of LMXBs is similar, that  $\sim 1000$  Galactic bulge LMXBs should have been detected, and therefore the detection of only 42 by the INTEGRAL telescope is evidence against the MSP scenario of the GCE. A similar argument is made by Cholis et al. [61]; however, these environments are different, and there are a number of pathways that could lead to recycled MSPs [62], the ratios of which may differ between globular clusters and the Galactic bulge. Bartels et al. [63] suggest that a population of dim ultra-compact X-ray binaries, a hydrogen-deficient sub-class of LMXBs that form at a significantly higher rate than hydrogen-rich LMXBs [64], could be consistent with the detected numbers and could produce enough MSPs to explain the GCE.

An alternative channel for MSP formation is Accretion Induced Collapse (AIC), in which a white dwarf accretes enough material to exceed the Chandrasekhar mass limit of about  $1.4M_\odot$  and collapses to a neutron star [54, 56, 65]. Conservation of angular momentum

could result in periods on the order of milliseconds, similarly magnetic flux conservation could produce  $\sim 10^8\text{--}10^9$  G magnetic field strengths after the collapse of a white dwarf where generally  $B \lesssim 10^3$  G [54]. We therefore do not need any process by which magnetic fields decay to explain observed MSP magnetic field strengths. Ferrario and Wickramasinghe [54] note that the distribution of magnetic field strengths for white dwarfs is bimodal, for the second group in the range  $\sim 10^6\text{--}10^9$  G AIC would produce MSPs with very strong magnetic fields  $\gtrsim 10^{10}$  G. However, they would spin down very rapidly. The kick associated with AIC would be much smaller than that occurring during core collapse of a massive star, this may explain the relatively low velocities of MSPs as well as the number in globular clusters, which have small escape velocities [65].

Neutron stars formed in core collapse supernovae receive a natal kick due to asymmetries in the supernova explosion, and these kick velocities are on the order of several hundred  $\text{km s}^{-1}$  [66, 67]. In Hobbs et al. [68], where the proper motions of 233 pulsars were studied, the young pulsars had mean transverse velocities of  $246 \pm 22 \text{ km s}^{-1}$ , however, the MSPs had significantly lower mean transverse velocities of  $87 \pm 13 \text{ km s}^{-1}$ . If MSPs are primarily recycled pulsars, then the lower velocities may be a result of the requirement that the binary system is not disrupted by a large natal kick. If it is disrupted, the neutron star may be observed as an isolated young pulsar with a high transverse velocity, but cannot later be spun up to millisecond periods by a binary companion. As part of the modelling of the Galactic MSP population in Chapter 2, we fit a Maxwell distributed peculiar velocity distribution to the resolved MSP proper motions, where peculiar velocity is the velocity relative to circular motion at a particular location in the Galaxy. We use the result as an input to the work described in Chapter 3, where we relate it to a natal kick velocity distribution.

The gamma-ray emission of MSPs is believed to be due to curvature radiation from charged particles accelerated in the electromagnetic field [18, 69]. Petri [69] uses a simple model in which a spherically symmetric distribution of positrons and electrons are assumed to be in equilibrium between acceleration in the electromagnetic field and braking due to the emission of radiation. In this radiation reaction limit, where the speed is assumed to be the speed of light  $c$ , the velocity of a positron or electron at a point depends on the local electric field  $\mathbf{E}$  and magnetic field  $\mathbf{B}$  [70]:

$$\mathbf{v}_{\pm} = \frac{\mathbf{E} \times \mathbf{B} \pm (E_0 \mathbf{E}/c + c B_0 \mathbf{B})}{E_0^2/c^2 + B^2} \quad (1.11)$$

where  $\mathbf{v}_+$  and  $\mathbf{v}_-$  are the velocity vectors for positrons and electrons respectively; and  $E_0$  and  $B_0$  are two electromagnetic invariants which, requiring  $E_0 \geq 0$ , satisfy:

$$E^2 - c^2 B^2 = E_0^2 - c^2 B_0^2 \quad (1.12)$$

$$\mathbf{E} \cdot \mathbf{B} = E_0 B_0 \quad (1.13)$$

Then the curvature radius  $\rho_c$  can be found using:

$$|\mathbf{a}_{\pm}| = \left| \frac{d\mathbf{v}_{\pm}}{dt} \right| = \frac{c^2}{\rho_c} \quad (1.14)$$

and the Lorentz factor  $\gamma$  is:

$$\gamma^4 = \frac{6\pi\epsilon_0}{e} E_0 \rho_c^2 \quad (1.15)$$

where  $\epsilon_0$  is the vacuum permittivity and  $e$  is the elementary charge. Petri [69] then uses for the curvature radiation the following spectrum:

$$\frac{dI}{d\omega} = \frac{\sqrt{3}e^2}{4\pi\epsilon_0 c} \gamma \frac{\omega}{\omega_c} \int_{\omega/\omega_c}^{\infty} K_{5/3}(x) dx \quad (1.16)$$

where  $I$  is the intensity,  $\omega$  is the angular frequency,  $K_{5/3}$  is the modified Bessel function of the second kind and where the characteristic frequency is:

$$\omega_c = \frac{3}{2} \gamma^3 \frac{c}{\rho_c} \quad (1.17)$$

In this simple model, with photons approximately emitted in the lepton direction of motion,  $\gamma \gtrsim 10^8$  producing pulsar spectra with peaks of a few GeV. If a particle trajectory is at a high pitch-angle relative to the asymptotic trajectory given by Eq. 1.11, the particle will emit synchrotron radiation where the radius of curvature used in Eqs. 1.16 and 1.17 is instead the gyro-radius of the particle motion. Kalapotharakos et al. [18] show that synchrotron radiation and curvature radiation will lead to different relationships between the luminosity of a pulsar and its spectral energy cutoff, magnetic field strength and spin down power. They show using a least squares fit to Fermi-LAT detected pulsars that the curvature radiation relationship provides a better fit to the data, and we confirm this in Chapter 2.

More realistic models of MSP emission focus on three regions, the polar caps, the slot gap and the outer gap [71–75]. The polar caps are a region where open magnetic field lines cross the surface of the pulsar. These field lines are those that cross the light cylinder at a radius  $r_L = cP/2\pi$ , beyond which particles cannot corotate with the neutron star with  $v < c$ . In the

polar cap model, the emission is associated with particles accelerated along open magnetic field lines from near the surface [76–79]. The slot gap is the region near the boundary of the closed field lines, extending from the surface of the pulsar to the light cylinder [80–82]. The outer gap is the volume in the light cylinder between closed field line boundary and the null surface where  $\mathbf{\Omega} \cdot \mathbf{B} = 0$ , where  $\mathbf{\Omega}$  is the rotation axis [83–86]. Kalapotharakos et al. [87] use a three dimensional particle-in-cell method to model the distribution and trajectories of charged particles in the pulsar magnetosphere and their associated magnetic fields. Using the radiation reaction limit, they simulate sky maps and spectra of the gamma-ray emission, they also show there is a relationship between spin down power and spectral energy cutoff.

# Chapter 2

## Modelling the Galactic Millisecond Pulsar Population

### 2.1 Introduction

In previous investigations, the spectrum of the MSPs in the bulge have been assumed to be the same as those in the disk. However, the bulge MSPs are expected to be on average older than the disk MSPs due to their different star formation histories [88]. Therefore, we would expect them to have on average lower luminosities. In addition to this, a correlation between luminosities and the spectral parameters is seen in the data [18, 45] and so one would expect the bulge MSPs to be spectrally different from the disk ones. This motivates the more detailed modelling of the MSP populations that is performed in the current chapter.

We extend the model of the Galactic population of gamma-ray MSPs of Ploeg et al. [15]. In that work, only distance estimates from parallax measurements were included. Here we also use the dispersion measure estimates of the MSP distances and incorporate the corresponding uncertainties in the free electron densities. Also, instead of empirically parameterising the MSP luminosity function, we start from empirical distributions describing MSP initial period, magnetic field strength, age, and gamma-ray spectra. We assume there is a relationship between the luminosity of a pulsar and some of its other properties such as its period, period derivative or spectral energy cutoff. This supposition is motivated by work such as by Kalapotharakos et al. [18] in which, based on data on resolved MSPs and young pulsars in Abdo et al. [45], it was determined that there is a relationship between MSP luminosity, on the one hand, and MSP spectral energy cutoff  $E_{\text{cut}}$ , magnetic field strength  $B$ , and the spin-down power  $\dot{E}$ , on the other. Specifically, Kalapotharakos et al. [18] found that gamma-ray emission from MSPs is via curvature radiation and scales

like  $L \propto E_{\text{cut}}^{1.18 \pm 0.24} B^{0.17 \pm 0.05} \dot{E}^{0.41 \pm 0.08}$  where we use 68% confidence intervals when quoting uncertainties unless otherwise specified. Similarly, Gonthier et al. [19] assumed that the radio and gamma ray luminosities of MSPs were dependent on period and period derivative and, under this assumption, successfully determined the parameters of that relationship using a model of the distribution and properties of the Galactic population of MSPs, radio and gamma-ray detection thresholds, and a model of how the observed flux of an individual MSP depends on viewing angle and magnetic axis angle.

## 2.2 Method

To fit our model of the Galactic MSP population we used MSPs with confirmed gamma ray pulsations according to the Public List of LAT-Detected Gamma-Ray Pulsars.<sup>1</sup> From that list, we obtained names and periods of pulsars. We then used the gamma-ray data for the corresponding sources in the Fermi Large Area Telescope fourth source catalog data release 2 [4FGL-DR2: 89]. Additional data were obtained from the ATNF pulsar catalog [52] if available. As in Bartels et al. [16] we used pulsars with periods less than 30 ms that were not associated with globular clusters. For the GCE, we use the boxy bulge and nuclear bulge spectra<sup>2</sup> from Macias et al. [8]. We included both the systematic and statistical errors of these spectra which were added in quadrature to get the total error. The systematic error accounts for variation in the GCE spectra caused by using different maps of Inverse Compton emission.

### 2.2.1 Modeling the Galactic Millisecond Pulsar population

#### Spatial distribution

The spatial model of MSPs has three components: a disk distribution, a boxy bulge distribution, and a nuclear bulge distribution. The disk component models the population from which we expect the resolved MSPs to mainly come and has density:

$$\rho_{\text{disk}}(R, z) \propto \exp(-R^2/2\sigma_r^2) \exp(-|z|/z_0) \quad (2.1)$$

where  $R^2 = x^2 + y^2$  is the radial coordinate in the Galactic disk and  $z$  is the height above the Galactic Plane. We treat  $\sigma_r$  and  $z_0$  as free parameters to be fit to the data. The modeled GCE

<sup>1</sup><https://confluence.slac.stanford.edu/display/GLAMCOG/Public+List+of+LAT-Detected+Gamma-Ray+Pulsars>

<sup>2</sup>Available from [https://github.com/chrisgordon1/Galactic\\_bulge\\_spectra](https://github.com/chrisgordon1/Galactic_bulge_spectra).



is produced by the boxy bulge and nuclear bulge components. The boxy bulge has density [1, 8]:

$$\rho_{\text{boxy bulge}}(R_s) \propto \text{sech}^2(R_s) \times \begin{cases} 1 & R \leq R_{\text{end}} \\ \exp(-(R - R_{\text{end}})^2/h_{\text{end}}^2) & R > R_{\text{end}} \end{cases} \quad (2.2)$$

where  $R_{\text{end}} = 3.128$  kpc,  $h_{\text{end}} = 0.461$  kpc, and:

$$R_{\perp}^{C_{\perp}} = \left( \frac{|x'|}{1.696 \text{ kpc}} \right)^{C_{\perp}} + \left( \frac{|y'|}{0.6426 \text{ kpc}} \right)^{C_{\perp}} \quad (2.3)$$

$$R_s^{C_{\parallel}} = R_{\perp}^{C_{\perp}} + \left( \frac{|z'|}{0.4425 \text{ kpc}} \right)^{C_{\parallel}} \quad (2.4)$$

where  $C_{\parallel} = 3.501$  and  $C_{\perp} = 1.574$ . The coordinates  $x'$ ,  $y'$  and  $z'$  are Cartesian coordinates in the boxy bulge frame. Relative to the frame in which  $x_{\odot} = -R_0$ ,  $y_{\odot} = z_{\odot} = 0$ , this frame is rotated  $13.79^\circ$  around the  $z$ -axis then  $0.023^\circ$  around the new  $y$ -axis. We assume  $R_0 = 8.3$  kpc as adopted from the YMW16 free electron density model [90] which we use to convert between distance and dispersion measure. The nuclear bulge MSP density is proportional to the sum of the mass densities of the NSC and NSD [7]:

$$\rho_{\text{nuclear bulge}}(r, z) \propto \rho_{\text{NSC}}(r) + \rho_{\text{NSD}}(r, z) \quad (2.5)$$

where  $r^2 = x^2 + y^2 + z^2$  and where the NSC has density:

$$\rho_{\text{NSC}}(r) = \begin{cases} \frac{\rho_{0,\text{NSC}}}{1 + \left(\frac{r}{r_0}\right)^2} & r \leq 6 \text{ pc} \\ \frac{\rho_{1,\text{NSC}}}{1 + \left(\frac{r}{r_0}\right)^3} & 6 \text{ pc} < r \leq 200 \text{ pc} \\ 0 & r > 200 \text{ pc} \end{cases} \quad (2.6)$$

where  $r_0 = 0.22$  pc,  $\rho_{0,\text{NSC}} = 3.3 \times 10^6 \text{ M}_{\odot} \text{ pc}^{-3}$  and  $\rho_{1,\text{NSC}}$  is set so that  $\rho_{\text{NSC}}(r)$  is continuous at  $r = 6$  pc. The NSD has density:

$$\rho_{\text{NSD}}(r, z) = \begin{cases} \rho_{0,\text{NSD}} \left( \frac{r}{1 \text{ pc}} \right)^{-0.1} e^{-\frac{|z|}{45 \text{ pc}}} & r < 120 \text{ pc} \\ \rho_{1,\text{NSD}} \left( \frac{r}{1 \text{ pc}} \right)^{-3.5} e^{-\frac{|z|}{45 \text{ pc}}} & 120 \text{ pc} \leq r < 220 \text{ pc} \\ \rho_{2,\text{NSD}} \left( \frac{r}{1 \text{ pc}} \right)^{-10} e^{-\frac{|z|}{45 \text{ pc}}} & r \geq 220 \text{ pc} \end{cases} \quad (2.7)$$

where  $\rho_{0,\text{NSD}} = 301 \text{ M}_\odot \text{ pc}^{-3}$  and  $\rho_{1,\text{NSD}}$  and  $\rho_{2,\text{NSD}}$  are set so that  $\rho_{\text{NSD}}(r, z)$  is continuous at both  $r = 120 \text{ pc}$  and  $r = 220 \text{ pc}$ .

### Age distribution

The MSP age distribution is dependent on two distributions: the SFR, and the delay time distribution (DTD). The delay time is the time between star and MSP formation. We assume for the DTD a five bin distribution, linearly spaced between 0 Gyr and the age of the universe.

For the disk and boxy bulge components of the MSP population, the SFR is [88]:

$$\text{SFR}(z) = \max(10^{Az^2+Bz+C} - D, 0) \quad (2.8)$$

where for the disk MSPs  $A = -4.06 \times 10^{-2}$ ,  $B = 0.331$ ,  $C = 0.338$  and  $D = 0.771$ . For the boxy bulge  $A = -2.62 \times 10^{-2}$ ,  $B = 0.384$ ,  $C = -8.42 \times 10^{-2}$  and  $D = 3.254$ . The relationship between cosmological time  $t$  and redshift  $z$  is [91]:

$$t(z) = \frac{9.778 \text{ Gyr}}{h} \int_0^{1/(1+z)} \frac{dx}{x \sqrt{\Omega_\Lambda + \Omega_K x^{-2} + \Omega_M x^{-3} + \Omega_R x^{-4}}} \quad (2.9)$$

where we have used  $h = 0.67$ ,  $\Omega_\Lambda = 0.68$ ,  $\Omega_M = 1 - \Omega_\Lambda$  and  $\Omega_K = \Omega_R = 0$  [92]. For the nuclear bulge, we use the MIST star formation rate from Nogueras-Lara et al. [93]. The three star formation rates are shown in Fig. 2.1.

From the SFR and DTD, we can find the probability density function of MSP ages:

$$p(\text{age}) = \frac{\int_0^{t_0 - \text{age}} \text{SFR}(z(\tau)) \text{DTD}(t_0 - \text{age} - \tau) d\tau}{\int_0^{t_0} \int_0^{t'} \text{SFR}(z(\tau)) \text{DTD}(t' - \tau) d\tau dt'} \quad (2.10)$$

where  $t_0$  is the age of the universe, and  $z(t)$  is the inverse of Eq. 2.9.

As an alternative age distribution, we try a uniform distribution where MSPs form at a constant rate over the last 10 Gyr, similar to that assumed by Gonthier et al. [19].

### Angular velocity

In our model, the angular transverse velocities of pulsars in the directions of Galactic longitude and latitude,  $\mu_l$  and  $\mu_b$  respectively, are determined by assuming a pulsar travels in a circular orbit around the center of the Galaxy using a parametric form of the potential [94, 95] with a random normally distributed peculiar velocity in every direction of scale  $\sigma_v$  (i.e., a Maxwell distribution). For the Sun, we assume circular motion in the same potential

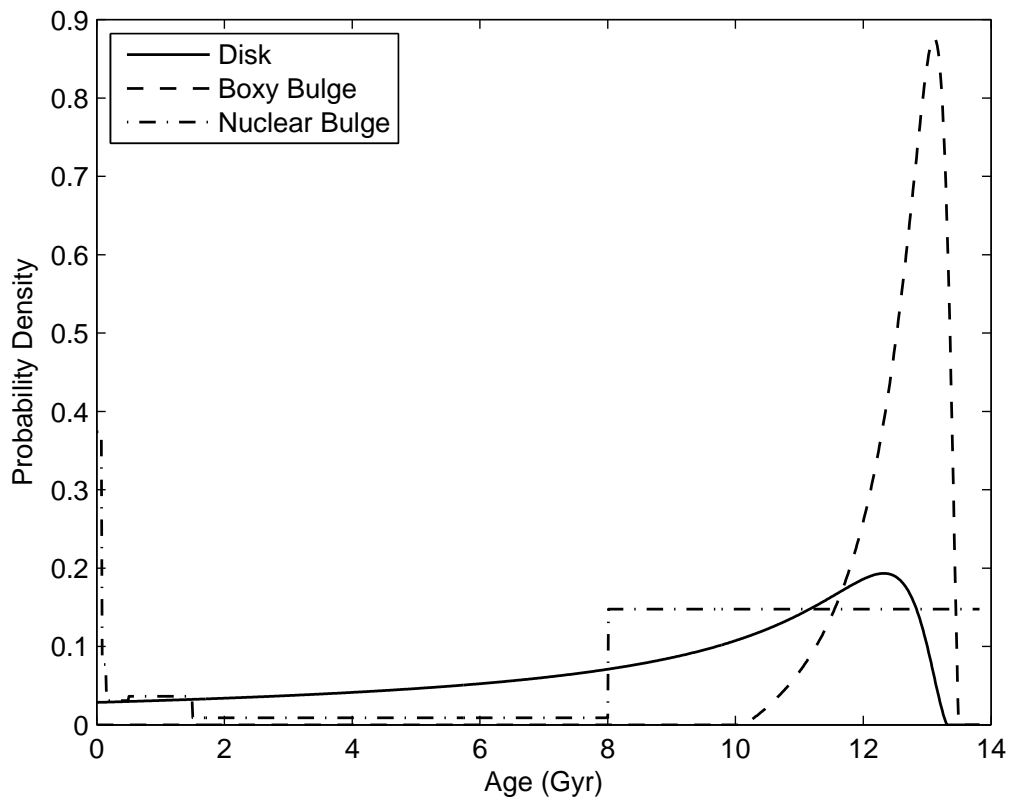


Fig. 2.1 Star formation rates for the disk, boxy bulge, and nuclear bulge. Here, this is given as a normalized probability density of stellar mass having formed at a particular age.

with a peculiar velocity of (11.1, 12.24, 7.25) km s<sup>-1</sup> where the velocity components are in the direction of the Galactic Center, the direction of rotation and in the direction perpendicular to the plane, respectively [96]. The relationship between velocity ( $v$ ) and angular velocity ( $\mu$ ) at a distance  $d$  is:

$$v \approx 4.74 \left( \frac{d}{\text{kpc}} \right) \left( \frac{\mu}{\text{mas yr}^{-1}} \right) \text{ km s}^{-1} \quad (2.11)$$

### Galactic Center Excess

To simulate the GCE, we need to assign each bulge MSP a spectrum that describes its photon number flux ( $N$ ) at energy  $E$ :

$$\frac{dN}{dE} = K E^{-\Gamma} \exp\left(- (E/E_{\text{cut}})^{2/3}\right) \quad (2.12)$$

where the spectral parameters are  $K$ ,  $\Gamma$ , and  $E_{\text{cut}}$ . The proportionality constant  $K$  is determined via:

$$F = \int_{0.1 \text{ GeV}}^{100 \text{ GeV}} E \frac{dN}{dE} dE \quad (2.13)$$

where  $F$  is the energy flux. The simulated boxy bulge and nuclear bulge GCE spectra are then the sum of all MSP spectra in each of those two bulge populations of MSPs. This spectrum is equivalent to the one used to fit the resolved MSPs in the 4FGL-DR2 catalog [89] and it will also be the one we use for our resolved MSPs.

### Millisecond pulsar parameters

Force-free electrodynamic solutions have given the following expression for the spin-down luminosity [97]

$$L_{\text{sd}} \sim \frac{\mu^2 \Omega^4}{c^3} (1 + \sin^2 \alpha) \quad (2.14)$$

where  $\mu$  is the magnetic dipole moment,  $\Omega$  is the rotational angular velocity,  $c$  is the speed of light, and  $\alpha$  is the angle between the rotation and magnetic field axes. The magnetic field strength ( $B$ ) at the magnetic pole of the star is related to  $\mu$  by [97]

$$\mu = \frac{BR^3}{2} \quad (2.15)$$

where  $R$  is the radius of the neutron star and we use  $R = 12$  km. The rotational kinetic energy of the neutron star is given by the standard formula for a rotating body

$$E = \frac{1}{2} I \Omega^2 \quad (2.16)$$

where  $I$  is the neutron star's moment of inertia and we use  $I = 1.7 \times 10^{45}$  g cm<sup>2</sup>. Therefore, the spin-down power satisfies

$$\dot{E} = 4\pi^2 I \dot{P}_{\text{int}} / P^3 \quad (2.17)$$

where  $\dot{P}_{\text{int}}$  is the time derivative of the intrinsic period, which may be different to the observed period derivative ( $\dot{P}$ ), and  $P = 2\pi/\Omega$  is the period. Equating the spin-down luminosity (Eq. 2.14) to the spin down power (Eq. 2.17) and using Eq. 2.15 yields the following expression for the magnetic field strength of an MSP

$$B^2 = \frac{c^3 I P \dot{P}_{\text{int}}}{\pi^2 R^6 (1 + \sin^2(\alpha))}. \quad (2.18)$$

The angle  $\alpha$  is chosen randomly from the probability density:

$$p(\alpha) = \frac{1}{2} \sin(\alpha) \quad (2.19)$$

which corresponds to a uniformly random magnetic field axis relative to the rotation axis.

The intrinsic period time derivative is related to the observed period time derivative  $\dot{P}$  by [45]:

$$\dot{P} = \dot{P}_{\text{int}} + \dot{P}_{\text{Shklovskii}} + \dot{P}_{\text{Galactic}} \quad (2.20)$$

where the contribution to the observed period derivative from the Shklovskii effect is given by

$$\dot{P}_{\text{Shklovskii}} = 2.43 \times 10^{-21} \left( \frac{\mu}{\text{mas yr}^{-1}} \right)^2 \left( \frac{d}{\text{kpc}} \right) \left( \frac{P}{\text{s}} \right) \quad (2.21)$$

and the contribution due to the relative acceleration in the Galactic potential is:

$$\dot{P}_{\text{Galactic}} = \frac{1}{c} \mathbf{n}_{10} \cdot (\mathbf{a}_p - \mathbf{a}_{\odot}) P \quad (2.22)$$

where  $\mathbf{n}_{10}$  is the unit vector from the Sun to the pulsar, and  $\mathbf{a}_p$  and  $\mathbf{a}_{\odot}$  are the accelerations, due to the Galactic potential [94, 95], of the pulsar and Sun respectively. Note that  $P = (1 + v_R/c) P_{\text{int}}$ , where  $v_R$  is the radial velocity of the pulsar and  $P_{\text{int}}$  is the intrinsic period. Given  $v_R \ll c$ , we approximate  $P_{\text{int}} = P$ . Making the common assumption that magnetic field

strength remains constant over time, Eq. 2.18 results in:

$$P = \sqrt{P_I^2 + \frac{2\pi^2 R^6}{I_C^3} (1 + \sin^2(\alpha)) B^2 t} \quad (2.23)$$

where  $P_I$  is the initial period of the MSP at birth and  $t$  is the age.

We consider multiple relationships between pulsar parameters and the 0.1-100 GeV luminosity ( $L$ ). Our most general form is that used by Kalapotharakos et al. [18]:

$$L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma} \quad (2.24)$$

where  $\eta$  is a proportionality factor. We also consider the model used by Gonthier et al. [19]:

$$L = \eta P^{\alpha_\gamma} \dot{P}^{\beta_\gamma}. \quad (2.25)$$

The simplest form we consider is that the luminosity is an entirely independent parameter as used by Ploeg et al. [15]:

$$L = \eta. \quad (2.26)$$

The likelihood probability density distributions of  $B$ ,  $E_{\text{cut}}$ , and  $\eta$  are assumed to be log-normal as this functional form gives a good fit to their histogrammed data of the resolved Fermi-LAT MSPs:

$$p(\log_{10}(x)|x_{\text{med}}, \sigma_x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{(\log_{10}(x) - \log_{10}(x_{\text{med}}))^2}{2\sigma_x^2}\right) \quad (2.27)$$

where  $x_{\text{med}}$  is the median of  $x$  and  $\sigma_x$  is the standard deviation of  $\log_{10}(x)$ . We also assume  $P_I$  has this form but our results are not sensitive to this assumption. In particular, we found that the cut off power law model used by Gonthier et al. [19] gives similar results. Note that, even though  $E_{\text{cut}}$  is obtained from a fit of an individual MSP's spectral data, for our purposes it is treated as a directly measured datum rather than a parameter to be estimated.

Fits to the MSP gamma-ray data [45] have uncovered correlations between the spectral parameters and  $\dot{E}$ . To allow for this to potentially be an intrinsic property of the MSPs, we parameterise the median of the  $E_{\text{cut}}$  likelihood as:

$$\log_{10}(E_{\text{cut,med}}/\text{MeV}) = a_{E_{\text{cut}}} \log_{10}(\dot{E}/(10^{34.5} \text{ erg s}^{-1})) + b_{E_{\text{cut}}} \quad (2.28)$$

where  $a_{E_{\text{cut}}}$  and  $b_{E_{\text{cut}}}$  are allowed to vary in our model fits. We also model the likelihood of spectral index  $\Gamma$  using a normal distribution with mean  $\mu_\Gamma$  and standard deviation  $\sigma_\Gamma$ . We

assume that

$$\mu_\Gamma = a_\Gamma \log_{10}(\dot{E}/(10^{34.5} \text{ erg s}^{-1})) + b_\Gamma \quad (2.29)$$

where  $a_\Gamma$  and  $b_\Gamma$  are parameters. We also have a correlation coefficient,  $r_{\Gamma, E_{\text{cut}}}$  between  $\Gamma$  and  $\log_{10}(E_{\text{cut}})$ , so the likelihood is:

$$p(\Gamma, \log_{10}(E_{\text{cut}})|\boldsymbol{\theta}) = \frac{1}{2\pi\sigma_\Gamma\sigma_{E_{\text{cut}}}\sqrt{1-r_{\Gamma, E_{\text{cut}}}^2}} \exp\left(-\frac{z_{\Gamma, E_{\text{cut}}}}{2(1-r_{\Gamma, E_{\text{cut}}}^2)}\right) \quad (2.30)$$

where:

$$z_{\Gamma, E_{\text{cut}}} = \frac{(\Gamma - \mu_\Gamma)^2}{\sigma_\Gamma^2} + \frac{(\log_{10}(E_{\text{cut}}) - \log_{10}(E_{\text{cut, med}}))^2}{\sigma_{E_{\text{cut}}}^2} - \frac{2r_{\Gamma, E_{\text{cut}}}(\Gamma - \mu_\Gamma)(\log_{10}(E_{\text{cut}}) - \log_{10}(E_{\text{cut, med}}))}{\sigma_\Gamma\sigma_{E_{\text{cut}}}}. \quad (2.31)$$

Note that here and in the rest of this chapter we will use  $\boldsymbol{\theta}$  to indicate the relevant model parameters. In this case they are  $\boldsymbol{\theta} = \{a_{E_{\text{cut}}}, b_{E_{\text{cut}}}, a_\Gamma, B_\Gamma, \sigma_{E_{\text{cut}}}, \sigma_\Gamma, r_{\Gamma, E_{\text{cut}}}\}$ . See Table 2.1 for the corresponding priors that we use.

As alternatives, we try two likelihoods of  $P_I$  where there is a dependence on  $B$ . In the first, a bivariate normal distribution relating  $\log_{10}(B)$  and  $\log_{10}(P_I)$  with a correlation parameter  $r_{B, P_I}$  is assumed. In the second, we adopt the relationship used by Gonthier et al. [19]:

$$P_I = 0.18 \times 10^{C_{P_I} + 3\delta/7} B_8^{6/7} \text{ ms} \quad (2.32)$$

where there is a lower bound of 1.3 ms,  $\delta$  is drawn from a uniform distribution between 0 and 2. Also,  $B_8 = (B/10^8 \text{ G})$ . We fit  $C_{P_I}$  as model parameter while it is set to 0 by Gonthier et al. [19].

### Millisecond pulsar detection

The flux of an MSP is related to the luminosity in the usual way:

$$F = \frac{L}{4\pi d^2} \quad (2.33)$$

and the MSP detection threshold flux  $F_{\text{th}}$  is drawn from a log-normal distribution so that the probability of a detection is [13, 15, 16]:

$$p(F_{\text{th}} \leq F|l, b, \boldsymbol{\theta}) = \frac{1}{2} \left( 1 + \text{erf} \left( \frac{\log_{10}(F) - (\log_{10}(\mu_{\text{th}}(l, b)) + K_{\text{th}})}{\sqrt{2}\sigma_{\text{th}}} \right) \right) \quad (2.34)$$

where  $K_{\text{th}}$  and  $\sigma_{\text{th}}$  are free parameters, and a map in longitude ( $l$ ) and latitude ( $b$ ) of  $\mu_{\text{th}}(l, b)$  associated with the 4FGL-DR2 catalog can be found online.<sup>3</sup>

### 2.2.2 Fitting the model to data

To fit the model parameters to the resolved MSP and GCE data, we use an adaptive Markov Chain Monte Carlo (MCMC) algorithm [98]. The Metropolis-Hastings algorithm [99, 100] is a method for generating a sequence of samples  $\mathbf{x}_0, \dots, \mathbf{x}_n$  that, for a sufficiently large number of samples, are distributed according to some probability distribution  $p(\mathbf{x})$ . Starting from some initial point  $\mathbf{x}_0$ , and for some proposal distribution  $q(\mathbf{y}|\mathbf{x})$ , we draw a sample  $\mathbf{x}_t$  as follows:

1. Draw a point  $\mathbf{y}$  from the proposal distribution  $q(\mathbf{y}|\mathbf{x}_{t-1})$
2. Calculate the acceptance probability  $r$ :

$$r = \frac{p(\mathbf{y})q(\mathbf{x}_{t-1}|\mathbf{y})}{p(\mathbf{x}_{t-1})q(\mathbf{y}|\mathbf{x}_{t-1})} \quad (2.35)$$

3. Set  $\mathbf{x}_t$  to  $\mathbf{y}$  with probability  $\min(1, r)$ , otherwise  $\mathbf{x}_t = \mathbf{x}_{t-1}$

The proposal distribution would typically be symmetric (such as a Gaussian centered at  $\mathbf{x}_{t-1}$ ) so that  $r$  simplifies to  $p(\mathbf{y})/p(\mathbf{x}_{t-1})$ . An inappropriate choice of proposal distribution could lead to a situation in which the number of samples required to converge to the target distribution  $p(\mathbf{x})$  could be unreasonably large. For example, a proposal distribution that generates too many small steps would take a long time to explore the parameter space. A proposal distribution that often produces steps where  $r$  is very small will take a long time to make a successful move.

The adaptive Metropolis algorithm of Haario et al. [98] is very similar to the Metropolis-Hastings algorithm except that the proposal distribution now depends on the history of the chain. We replace  $q(\mathbf{y}|\mathbf{x}_{t-1})$  with  $q(\mathbf{y}|\mathbf{x}_{t-1}, \dots, \mathbf{x}_0)$ . This proposal distribution is a Gaussian centered at  $\mathbf{x}_{t-1}$  with a covariance matrix  $C_t$  that adapts so that the proposal distribution is of a scale similar to the target distribution. For the first  $t_0$  samples of the chain we use some initial covariance matrix  $C_0$ , then use  $C_t$  related to the sample covariance matrix of all past samples  $\mathbf{x}_0, \dots, \mathbf{x}_{t-1}$ :

$$C_t = \begin{cases} C_0 & t \leq t_0 \\ s_d \text{cov}(\mathbf{x}_0, \dots, \mathbf{x}_{t-1}) + s_d \epsilon I & t > t_0 \end{cases} \quad (2.36)$$

---

<sup>3</sup>[https://fermi.gsfc.nasa.gov/ssc/data/access/lat/10yr\\_catalog/](https://fermi.gsfc.nasa.gov/ssc/data/access/lat/10yr_catalog/)



where for  $d$  dimensions  $s_d = 2.4^2/d$  and the extra term  $s_d \varepsilon I$  with small  $\varepsilon$  prevents the covariance matrix from becoming singular.

The resolved MSPs have an unbinned Poisson likelihood [101]:

$$\mathcal{L}_{\text{res}} \propto \exp(-\lambda_{\text{res}}) \prod_{i=1}^{N_{\text{res}}} \rho(\mathbf{D}_i) \quad (2.37)$$

where  $N_{\text{res}}$  is the number of resolved MSPs,  $\lambda_{\text{res}}$  is the expected number of resolved MSPs, and  $\rho(\mathbf{D}_i)$  is the phase space density of the resolved MSPs which have the data

$$\mathbf{D}_i = \{l_i, b_i, d_i, P_i, \dot{P}_i, \mu_{l,i}, \mu_{b,i}, F_i, E_{\text{cut},i}, \Gamma_i\}. \quad (2.38)$$

That is, the resolved MSP is located at longitude  $l_i$ , latitude  $b_i$ , distance  $d_i$ , period  $P_i$ , observed period derivative  $\dot{P}_i$ , proper motion in longitude  $\mu_{l,i}$ , proper motion in latitude  $\mu_{b,i}$ , flux  $F_i$ , spectrum energy cut-off  $E_{\text{cut},i}$ , and spectral index  $\Gamma_i$ .

The expected number of resolved MSPs can be obtained by integrating the phase space density over the phase space volume as follows [101]:

$$\lambda_{\text{res}} = \int \rho(\mathbf{D}_i) d\mathbf{D}_i \quad (2.39)$$

This relation implies that

$$\rho(\mathbf{D}_i) = p(\mathbf{D}_i|\text{obs}, \boldsymbol{\theta}) \lambda_{\text{res}} \quad (2.40)$$

where obs indicates the MSP was observed, i.e. it was resolved. Also,  $p(\mathbf{D}_i|\text{obs}, \boldsymbol{\theta})$  is the probability density that a resolved MSP has data  $\mathbf{D}_i$  given that the model parameter values are  $\boldsymbol{\theta}$ . It then follows from the above two equations that

$$\int p(\mathbf{D}_i|\text{obs}, \boldsymbol{\theta}) d\mathbf{D}_i = 1 \quad (2.41)$$

as required.

There are many tens of thousands of unresolved MSPs in the Milky Way [15, 19] and we have only resolved of order 100 in gamma rays. It follows that the probability of observing an individual MSP must be a very small. Therefore, from the law of rare events (see for example Section 1.1.1 of [102]), the total number of resolved and unresolved MSPs ( $N_{\text{tot}}$ ) is well approximated by:

$$N_{\text{tot}} = \lambda_{\text{res}} / p(\text{obs}|\boldsymbol{\theta}). \quad (2.42)$$

We can find  $p(\text{obs}|\boldsymbol{\theta})$  by noting that a luminosity threshold distribution  $p(L_{\text{th}})$  is determined by the combination of the flux threshold and spatial distributions. Thus:

$$p(\text{obs}|\boldsymbol{\theta}) = \int dL_{\text{th}} p(L_{\text{th}}) p(L \geq L_{\text{th}}) \quad (2.43)$$

where  $p(L \geq L_{\text{th}})$  is the probability that the luminosity is greater than or equal to the threshold. Another useful relation that follows from standard probability theory is:

$$p(\text{obs}, \mathbf{D}_i|\boldsymbol{\theta}) = p(\mathbf{D}_i|\text{obs}, \boldsymbol{\theta}) p(\text{obs}|\boldsymbol{\theta}). \quad (2.44)$$

Combining the above equation with Eqs. 2.42 and 2.40 gives

$$\rho(\mathbf{D}_i) = p(\text{obs}, \mathbf{D}_i|\boldsymbol{\theta}) N_{\text{tot}}. \quad (2.45)$$

We have two distinct types of resolved MSPs: those with parallax measurements and those without. To accommodate this we have two separate probability density functions

$$p(\text{obs}, \mathbf{D}_i|\boldsymbol{\theta}) = p(\text{obs}, \mathbf{D}_i, \text{parallax}_i|\boldsymbol{\theta}) + p(\text{obs}, \mathbf{D}_i, \text{not parallax}_i|\boldsymbol{\theta}) \quad (2.46)$$

The  $\mathbf{D}_i$  components are the same in both cases as we can estimate the distance for those MSPs that do not have parallax measurements by their dispersion measures. The probability of a parallax measurement given distance  $d$  is modelled as:

$$p(\text{parallax}_i|d_i, \boldsymbol{\theta}) = \min(1, C_{\text{parallax}} \exp(-d_i/d_{\text{parallax}})). \quad (2.47)$$

and  $p(\text{not parallax}_i|d_i, \boldsymbol{\theta}) = 1 - p(\text{parallax}_i|d_i, \boldsymbol{\theta})$ .

To take into account measurement uncertainty in our values of  $\mathbf{D}_i$  we marginalise over the true values  $\hat{\mathbf{D}}_i$

$$p(\text{obs}, \mathbf{D}_i|\boldsymbol{\theta}) = \int p(\mathbf{D}_i|\hat{\mathbf{D}}_i) p(\text{obs}, \hat{\mathbf{D}}_i|\boldsymbol{\theta}) d\hat{\mathbf{D}}_i \quad (2.48)$$

If MSP  $i$  has the  $j$ th component of its data  $\mathbf{D}_i$  missing then we account for that by making  $p(\mathbf{D}_i|\hat{\mathbf{D}}_i)$  uniform in the  $j$ th component.

Some of the data used are given in the form of  $\mathbf{y}_j$  where  $\mathbf{y}$  is not one of the directly modeled MSP parameters. In this case we perform a transformation to the observed parameter:

$$p(\dots, \mathbf{y}_j, \dots) = \int d\mathbf{y} p(\mathbf{y}_j|\mathbf{y}) |J| p(\dots, \mathbf{x}(\mathbf{y}), \dots) \quad (2.49)$$

where  $|\mathbf{J}|$  is the Jacobian determinant of the transformation from  $\mathbf{y}$  to  $\mathbf{x}$ , so  $\mathbf{J}_{i,j} = \partial x_i(\mathbf{y}) / \partial y_j$ . This transformation is needed for integrating over uncertainty in distance where we have either parallax ( $\omega$ ), where  $\omega = 1/d$ , or dispersion measure (DM), where  $\text{DM} = \int_0^d n_e(s) ds$  with  $n_e(s)$  being the free electron density at distance  $s$ .

See Appendix A for more details of the resolved MSP probability density function and Appendix B for more details about the measurement uncertainties.

We fit the boxy bulge and nuclear bulge GCE to spectra found by Macias et al. [8]. We use a Gaussian likelihood for each bin in the GCE spectra:

$$\mathcal{L}_{\text{GCE}} \propto \prod_{i=1}^N \exp \left( - \left( \left( \frac{dN}{dE} \right)_{\text{sim},i} - \left( \frac{dN}{dE} \right)_{\text{data},i} \right)^2 / (2\sigma_{\text{data},i}^2) \right) \quad (2.50)$$

where  $(\frac{dN}{dE})_{\text{sim},i}$  is the simulated GCE for bin  $i$  and  $(\frac{dN}{dE})_{\text{data},i}$  is the data with uncertainty  $\sigma_{\text{data},i}^2$ . We fit only energy bins lower than 10 GeV as the higher energy bins may contain significant secondary emission [8]. The combined likelihood for our resolved MSPs and the GCE is then given by substituting Eqs. 2.50 and 2.37 into

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{res}} \times \mathcal{L}_{\text{GCE}}. \quad (2.51)$$

This is then multiplied with the priors given in Table 2.1 to get the posterior which is then sampled using MCMC:

$$p(\boldsymbol{\theta} | \mathbf{D}_0, \dots, \mathbf{D}_{N_{\text{res}}}) \propto p(\boldsymbol{\theta}) \mathcal{L}_{\text{total}} \quad (2.52)$$

where  $p(\boldsymbol{\theta})$  is the prior probability density function. To test our models we generate posterior predictive distributions [100]. See Appendix C for more details on the sampling methods we used.

## 2.3 Results

In this section we present results for various assumed luminosity functions, age distributions, and relationships between magnetic field strength and initial period. We rank these various models using the Watanabe Akaike Information Criterion (WAIC) [100, 103] as described in Appendix D. The WAIC provides a measure of the expected predictive accuracy of a model and it takes into account the number of model parameters and their posterior uncertainty. Under suitable regularity conditions, in the limit of a large amount of data, the difference in the WAIC between two models ( $\Delta\text{WAIC}$ ) tends towards minus two times the log of their

likelihood ratio [100]. So when adding an extra parameter to a model, the “number of sigma” in favour of adding that parameter is given approximately by  $\sqrt{\Delta\text{WAIC}}$  [104]. This provides a rough benchmark in evaluating the significance of  $\Delta\text{WAIC}$  values.

We can gain a more detailed view of a model’s fit by comparing its posterior predictive distributions to data using single dimensional binned plots showing medians and 68% and 95% intervals, as well as corner plots constructed using the software by Foreman-Mackey [105]. These posterior predictive distributions provide an effective method of evaluating the goodness of fit [100, 106]. In particular, major failures of the model correspond to extreme posterior predictive p-values. These are defined as the proportion of posterior simulations which are more extreme than the data or some statistic of the data [100]. Our corner plots show two dimensional distributions of real data and simulated data with 68%, 95% and 99.7% contours. In producing simulated data we model both missing data and uncertainties by picking a random real MSP and removing simulated data that are not available for the real MSP. For data with uncertainties attached, we take the relative error for the real MSP and add Gaussian noise to the simulated MSP which has a standard deviation with the same relative error as the real MSP. We have a distance dependent model of the probability of a parallax measurement being available. If a simulated MSP has a parallax measurement, with probability given by Eq. 2.47, we select a random real MSP out of those with available parallax measurements and use its relative uncertainty to simulate a parallax error.

The WAIC allows us to compare the various models used while accounting for the varying number of parameters they involve. In Table 2.2 we show the WAIC averaged over the eight chains for each of a set of models of the Galactic MSP population relative to that of the best model. We also show the sample standard deviation over the WAIC for the eight MCMC chains run for each model. This variation occurs because we used Monte Carlo integration to compute the integrals in Section 2.2. The random numbers used to compute these integrals (such as over the resolved MSP data uncertainty distributions) were generated once per MCMC chain so that a calculation of the likelihood for a given set of parameters will always return the same result. However, between chains the likelihood may shift slightly as the set of randomly generated numbers used will be different. As a result of this variation in the likelihood, we did not generate a single WAIC for all eight chains combined. We found that the posterior distributions of the model parameters were generally indistinguishable despite these variations, and the variation in WAIC for each model is typically small compared to  $\Delta\text{WAIC}$  between models.

In Table 2.3 we report medians and 68% confidence intervals for the parameters of a subset of these models. The first three parameters in this table are related to the number of MSPs in each of the three spatial components of the model: The parameter  $\lambda_{\text{res}}$  is the

expected number of resolved MSPs;  $\log_{10}(N_{\text{disk}}/N_{\text{bulge}})$  and  $\log_{10}(N_{\text{nb}}/N_{\text{bb}})$  are parameters defining, respectively, the ratio of  $N_{\text{disk}}$  to  $N_{\text{bulge}}$  and  $N_{\text{nb}}$  to  $N_{\text{bb}}$ . Here  $N_{\text{disk}}$  is the number of disk MSPs,  $N_{\text{bulge}} = N_{\text{nb}} + N_{\text{bb}}$  is the total number of bulge MSPs,  $N_{\text{nb}}$  is the number of nuclear bulge MSPs, and  $N_{\text{bb}}$  is the number of boxy bulge MSPs.

In Figs. 2.2, 2.3, and 2.4, we display corner plots showing some of these parameters for the best model, which was A1 which had  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ . In Fig. 2.5 the resolved MSP and GCE data are compared to simulated data. We exclude data that led to an apparently negative  $\dot{P}_{\text{int}}$  from the binned  $\dot{P}$  data. This exclusion affects all bins with data except the highest two. The luminosity distribution is shown in Fig. 2.6, and the MSP age distributions are shown in Fig. 2.7. In Fig. 2.8 we compare the total gamma ray emission from MSPs in the region of interest to the observed total. In Fig. 2.9 we show the number of MSPs in the disk and in the bulge. The number with luminosity greater than  $10^{32} \text{ erg s}^{-1}$  is shown in Fig. 2.10, and in Fig. 2.11 we show the number of MSPs produced per solar mass at  $t = \infty$  assuming no further star formation after today. The masses used were  $(3.7 \pm 0.5) \times 10^{10} M_\odot$  for the disk,  $(1.6 \pm 0.2) \times 10^{10} M_\odot$  for the boxy bulge, and  $(1.4 \pm 0.6) \times 10^9 M_\odot$  for the nuclear bulge [88, 107]. In Fig. 2.12 we show the modeled probability of resolving  $N$  bulge MSPs, as well as the number for double and quadruple the current sensitivity. In Fig. 2.13 we compare the posterior predictive distributions for the boxy bulge spectra in the cases where we fitted the DTD and where the MSPs were uniformly distributed in age.

Finally, we also display corner plots showing the distributions of simulated, resolved MSPs for different models of the MSP luminosity function: Fig. 2.14 shows the  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$  case, Fig. 2.15 shows the  $L = \eta$  case, and Fig. 2.16 shows the case for the  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$  with  $E_{\text{cut}}$  and  $\Gamma$  independent of  $\dot{E}$ .

Parameter	Prior Minimum	Prior Maximum
$\lambda_{\text{res}}$	0	1000
$\log_{10}(N_{\text{disk}}/N_{\text{bulge}})$	-100	100
$\log_{10}(N_{\text{nb}}/N_{\text{bb}})$	-100	100
$\sigma_r$ (kpc)	0	15
$z_0$ (kpc)	0	1.5
$K_{\text{th}}$	-20	20
$\sigma_{\text{th}}$	0	5
$C_{\text{parallax}}$	0	5
$\log_{10}(d_{\text{parallax}}/\text{kpc})$	-1	2
$a_{E_{\text{cut}}}$	-5	5
$b_{E_{\text{cut}}}$	2	5
$\sigma_{E_{\text{cut}}}$	0	3
$a_{\Gamma}$	-5	5
$b_{\Gamma}$	0	5
$\sigma_{\Gamma}$	0	5
$r_{\Gamma, E_{\text{cut}}}$	-1	1
$a_{\gamma}$	-10	10
$b_{\gamma}$	-10	10
$d_{\gamma}$	-10	10
$\alpha_{\gamma}$	-10	10
$\beta_{\gamma}$	-10	10
DTD $p(0 - 2.8 \text{ Gyr})$	0	1
DTD $p(2.8 - 5.5 \text{ Gyr})$	0	1
DTD $p(5.5 - 8.3 \text{ Gyr})$	0	1
DTD $p(8.3 - 11.1 \text{ Gyr})$	0	1
DTD $p(11.1 - 13.8 \text{ Gyr})$	0	1
$\log_{10}(P_{l, \text{med}}/\text{s})$	-4	10
$\sigma_{P_l}$	0	10
$C_{P_l}$	-2	5
$\log_{10}(B_{\text{med}}/\text{G})$	-1	20
$\sigma_B$	0	10
$r_{B, P_l}$	-1	1
$\log_{10}(\eta_{\text{med}})$	-50	50
$\sigma_{\eta}$	0	10
$\sigma_v$ ( $\text{km s}^{-1}$ )	0	10000

Table 2.1 Prior ranges for model parameters. The priors are uniform within these ranges except for the DTD bin probabilities which have a Dirichlet prior [108]. I.e. the prior for the DTD bin probabilities is uniform on a four dimensional hyperplane in the five dimensional DTD bin space. The hyperplane consist of all points for which the five DTD bin probabilities add up to one.

Model Label	Description	$\Delta\text{WAIC}$	WAIC Std. Dev.	$p_{\text{WAIC}}$
A1	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$	0	0.8	26.9
A2	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , Uniform Age Distribution	0.4	1.0	26.1
A3	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , Covariance $B$ and $P_I$	1.5	1.0	27.5
A4	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , Eq. 2.32 Initial Period Distribution	7.3	1.5	25.9
A5	$L = \eta \dot{E}^{\alpha_\gamma}$	9.5	1.2	24.4
A6	$L = \eta P^{\alpha_\gamma} \dot{P}^{\beta_\gamma}$	11.0	0.7	25.3
A7	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , $a_{E_{\text{cut}}} = a_\Gamma = 0$	19.4	0.5	25.1
A8	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , $a_{E_{\text{cut}}} = a_\Gamma = 0$ , Uniform Age Distribution	21.1	0.7	24.0
A9	$L = \eta$	42.2	0.7	24.1
B1	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , Uniform Age Distribution, No GCE	0	1	24
B2	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , No GCE	1.2	0.8	25.3
B3	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , Covariance $B$ and $P_I$ , No GCE	2.3	0.8	26.1
B4	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , Eq. 2.32 Initial Period Distribution, No GCE	9	4	24
B5	$L = \eta \dot{E}^{\alpha_\gamma}$ , No GCE	13	1	23
B6	$L = \eta P^{\alpha_\gamma} \dot{P}^{\beta_\gamma}$ , No GCE	15.2	0.9	23.8
B7	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , $a_{E_{\text{cut}}} = a_\Gamma = 0$ , No GCE	21.3	0.9	23.6
B8	$L = \eta E_{\text{cut}}^{\alpha_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , $a_{E_{\text{cut}}} = a_\Gamma = 0$ , Uniform Age Distribution, No GCE	22.9	0.9	22.5
B9	$L = \eta$ , No GCE	47	2	23

Table 2.2 Average WAIC for each model relative to the model with best (lowest) WAIC. The average is taken over eight MCMC chains run for each model. We separate cases where we do not fit a GCE (B1-B9) from those where we do (A1-A9). The WAIC standard deviation column shows the sample standard deviation of WAIC for the eight MCMC chains run for each model. The  $p_{\text{WAIC}}$  column gives a measure of the effective number of parameters for the corresponding model [100, 103].

Parameter	Model A1	Model B2	Model A5	Model A6	Model A7	Model A9
$\lambda_{\text{res}}$	$108^{+11}_{-10}$	$107^{+11}_{-10}$	$108^{+10}_{-10}$	$108^{+10}_{-10}$	$108^{+11}_{-10}$	$108^{+11}_{-10}$
$\log_{10}(N_{\text{disk}}/N_{\text{bulge}})$	$0.00^{+0.13}_{-0.12}$	—	$0.00^{+0.13}_{-0.12}$	$0.01^{+0.13}_{-0.12}$	$0.04^{+0.12}_{-0.11}$	$0.09^{+0.11}_{-0.10}$
$\log_{10}(N_{\text{nb}}/N_{\text{bb}})$	$-0.66^{+0.08}_{-0.07}$	—	$-0.66^{+0.08}_{-0.07}$	$-0.66^{+0.07}_{-0.07}$	$-0.61^{+0.06}_{-0.07}$	$-0.58^{+0.04}_{-0.04}$
$\sigma_r$ (kpc)	$4.5^{+0.5}_{-0.4}$	$4.5^{+0.5}_{-0.4}$	$4.4^{+0.5}_{-0.4}$	$4.5^{+0.5}_{-0.4}$	$4.4^{+0.5}_{-0.4}$	$4.5^{+0.6}_{-0.4}$
$z_0$ (kpc)	$0.71^{+0.11}_{-0.09}$	$0.70^{+0.10}_{-0.09}$	$0.71^{+0.11}_{-0.09}$	$0.72^{+0.11}_{-0.09}$	$0.71^{+0.11}_{-0.09}$	$0.72^{+0.11}_{-0.09}$
$K_{\text{th}}$	$0.45^{+0.09}_{-0.08}$	$0.43^{+0.09}_{-0.08}$	$0.45^{+0.09}_{-0.08}$	$0.45^{+0.09}_{-0.08}$	$0.46^{+0.09}_{-0.08}$	$0.46^{+0.09}_{-0.08}$
$\sigma_{\text{th}}$	$0.28^{+0.05}_{-0.04}$	$0.27^{+0.05}_{-0.04}$	$0.28^{+0.05}_{-0.04}$	$0.28^{+0.05}_{-0.04}$	$0.28^{+0.05}_{-0.04}$	$0.28^{+0.05}_{-0.04}$
$C_{\text{parallax}}$	$0.43^{+0.15}_{-0.10}$	$0.45^{+0.16}_{-0.11}$	$0.43^{+0.15}_{-0.10}$	$0.42^{+0.15}_{-0.09}$	$0.43^{+0.15}_{-0.10}$	$0.40^{+0.13}_{-0.08}$
$\log_{10}(d_{\text{parallax}}/\text{kpc})$	$0.8^{+0.6}_{-0.3}$	$0.7^{+0.6}_{-0.3}$	$0.8^{+0.6}_{-0.3}$	$0.8^{+0.6}_{-0.3}$	$0.8^{+0.6}_{-0.3}$	$0.9^{+0.6}_{-0.4}$
$a_{E_{\text{cut}}}$	$0.18^{+0.05}_{-0.05}$	$0.21^{+0.05}_{-0.05}$	$0.19^{+0.05}_{-0.05}$	$0.19^{+0.05}_{-0.05}$	—	$0.17^{+0.05}_{-0.05}$
$b_{E_{\text{cut}}}$	$2.83^{+0.05}_{-0.06}$	$2.79^{+0.06}_{-0.07}$	$2.99^{+0.03}_{-0.02}$	$2.99^{+0.03}_{-0.02}$	$2.79^{+0.06}_{-0.06}$	$2.99^{+0.03}_{-0.03}$
$\sigma_{E_{\text{cut}}}$	$0.23^{+0.02}_{-0.02}$	$0.23^{+0.02}_{-0.02}$	$0.22^{+0.02}_{-0.02}$	$0.22^{+0.02}_{-0.02}$	$0.25^{+0.02}_{-0.02}$	$0.23^{+0.02}_{-0.02}$
$a_{\Gamma}$	$0.41^{+0.08}_{-0.08}$	$0.43^{+0.08}_{-0.08}$	$0.43^{+0.07}_{-0.08}$	$0.43^{+0.07}_{-0.08}$	—	$0.39^{+0.08}_{-0.08}$
$b_{\Gamma}$	$0.81^{+0.07}_{-0.08}$	$0.77^{+0.08}_{-0.10}$	$1.00^{+0.04}_{-0.04}$	$1.00^{+0.04}_{-0.04}$	$0.70^{+0.09}_{-0.10}$	$0.99^{+0.04}_{-0.04}$
$\sigma_{\Gamma}$	$0.36^{+0.04}_{-0.03}$	$0.36^{+0.04}_{-0.03}$	$0.34^{+0.03}_{-0.03}$	$0.34^{+0.03}_{-0.03}$	$0.42^{+0.04}_{-0.03}$	$0.35^{+0.03}_{-0.03}$
$r_{\Gamma, E_{\text{cut}}}$	$0.75^{+0.05}_{-0.06}$	$0.77^{+0.05}_{-0.06}$	$0.73^{+0.06}_{-0.07}$	$0.73^{+0.06}_{-0.07}$	$0.79^{+0.04}_{-0.05}$	$0.75^{+0.05}_{-0.06}$
$a_{\gamma}$	$1.2^{+0.3}_{-0.3}$	$1.4^{+0.3}_{-0.3}$	—	—	$1.3^{+0.3}_{-0.3}$	—
$b_{\gamma}$	$0.1^{+0.4}_{-0.4}$	$0.1^{+0.4}_{-0.4}$	—	—	$0.1^{+0.4}_{-0.4}$	—
$d_{\gamma}$	$0.50^{+0.12}_{-0.12}$	$0.48^{+0.12}_{-0.12}$	—	—	$0.56^{+0.11}_{-0.10}$	—
$\alpha_{\gamma}$	—	—	$0.74^{+0.11}_{-0.10}$	$-2.2^{+0.4}_{-0.4}$	—	—
$\beta_{\gamma}$	—	—	—	$0.79^{+0.20}_{-0.20}$	—	—
DTD $p(0 - 2.8 \text{ Gyr})$	$0.13^{+0.14}_{-0.09}$	$0.13^{+0.15}_{-0.09}$	$0.13^{+0.14}_{-0.09}$	$0.13^{+0.15}_{-0.09}$	$0.10^{+0.11}_{-0.07}$	$0.02^{+0.02}_{-0.01}$
DTD $p(2.8 - 5.5 \text{ Gyr})$	$0.15^{+0.18}_{-0.11}$	$0.15^{+0.19}_{-0.11}$	$0.15^{+0.18}_{-0.11}$	$0.15^{+0.19}_{-0.11}$	$0.12^{+0.16}_{-0.09}$	$0.02^{+0.03}_{-0.02}$



Parameter	Model A1	Model B2	Model A5	Model A6	Model A7	Model A9
DTD $p(5.5 - 8.3 \text{ Gyr})$	$0.29^{+0.20}_{-0.17}$	$0.26^{+0.18}_{-0.16}$	$0.29^{+0.20}_{-0.17}$	$0.29^{+0.20}_{-0.17}$	$0.25^{+0.18}_{-0.14}$	$0.09^{+0.06}_{-0.05}$
DTD $p(8.3 - 11.1 \text{ Gyr})$	$0.14^{+0.16}_{-0.10}$	$0.11^{+0.14}_{-0.08}$	$0.15^{+0.17}_{-0.10}$	$0.14^{+0.16}_{-0.10}$	$0.14^{+0.17}_{-0.10}$	$0.06^{+0.11}_{-0.05}$
DTD $p(11.1 - 13.8 \text{ Gyr})$	$0.14^{+0.22}_{-0.11}$	$0.22^{+0.20}_{-0.15}$	$0.13^{+0.21}_{-0.10}$	$0.14^{+0.21}_{-0.11}$	$0.27^{+0.24}_{-0.19}$	$0.78^{+0.08}_{-0.14}$
$\log_{10}(P_{i, \text{ med}}/\text{s})$	$-2.61^{+0.05}_{-0.04}$	$-2.60^{+0.05}_{-0.04}$	$-2.61^{+0.05}_{-0.04}$	$-2.61^{+0.05}_{-0.04}$	$-2.63^{+0.04}_{-0.03}$	$-2.69^{+0.03}_{-0.03}$
$\sigma_{P_i}$	$0.13^{+0.02}_{-0.02}$	$0.13^{+0.02}_{-0.02}$	$0.13^{+0.02}_{-0.02}$	$0.12^{+0.02}_{-0.02}$	$0.12^{+0.02}_{-0.02}$	$0.13^{+0.02}_{-0.02}$
$\log_{10}(B_{\text{med}}/\text{G})$	$8.21^{+0.05}_{-0.06}$	$8.21^{+0.06}_{-0.06}$	$8.21^{+0.03}_{-0.04}$	$8.20^{+0.06}_{-0.06}$	$8.21^{+0.05}_{-0.05}$	$8.25^{+0.02}_{-0.02}$
$\sigma_B$	$0.21^{+0.03}_{-0.02}$	$0.21^{+0.03}_{-0.02}$	$0.22^{+0.03}_{-0.03}$	$0.22^{+0.03}_{-0.02}$	$0.21^{+0.03}_{-0.03}$	$0.19^{+0.02}_{-0.02}$
$\log_{10}(\eta_{\text{med}})$	$12^{+5}_{-5}$	$12^{+5}_{-5}$	$7^{+4}_{-4}$	$-4^{+4}_{-4}$	$9^{+4}_{-4}$	$32.17^{+0.23}_{-0.31}$
$\sigma_\eta$	$0.52^{+0.06}_{-0.05}$	$0.53^{+0.06}_{-0.05}$	$0.58^{+0.06}_{-0.05}$	$0.58^{+0.07}_{-0.06}$	$0.51^{+0.06}_{-0.05}$	$0.72^{+0.08}_{-0.06}$
$\sigma_v (\text{km s}^{-1})$	$77^{+6}_{-6}$	$76^{+6}_{-5}$	$77^{+6}_{-6}$	$77^{+6}_{-6}$	$77^{+6}_{-6}$	$78^{+6}_{-6}$

Table 2.3 Medians and 68% confidence intervals for a selection of different models of the Galactic MSP population. See Table 2.2 for a description of the model labels.

## 2.4 Discussion

Some of our luminosity functions are nested. We can go from Model A1 with  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$  to Model A9 with  $L = \eta$  by setting  $a_\gamma = b_\gamma = d_\gamma = 0$ . As can be seen from the Model A1 fit in Fig. 2.2 and Table 2.3, the data prefer the  $a_\gamma = 1.2 \pm 0.3$  and  $d_\gamma = 0.5 \pm 0.1$  parameters to be larger than zero at high significance. This is consistent with the results of Table 2.2 where Model A9 has a  $\Delta\text{WAIC} = 42.2$  relative to Model A1. Comparing the posterior predictive corner plots for Model A1 in Fig. 2.14 and Model A9 in Fig. 2.15 we can see that Model A9 does not capture some of the correlations in the data, particularly those between distance (shown indirectly in the form of dispersion measure) and the spectral parameter  $E_{\text{cut}}$  as well as between distance and period. A comparison of correlation coefficients between the real data and simulated data is shown in Fig. 2.17. More detail can be seen by comparing Tables 2.4 and 2.7 which list the posterior predictive p-values for the correlations between observables for Models A1 and A9 respectively. Following the recommendation given in Section 6.3 of ref. [100], we consider posterior predictive p-values below 1% and above 99% to be of concern. As can be seen, the posterior predictive p-values for Model A1 are all comfortably within the 1% to 99% interval. In contrast to this, seven correlation coefficients for Model A9 are outside this interval. The correlation coefficient between  $\log_{10}(P)$  and  $\log_{10}(\text{DM})$  and also the one between  $\log_{10}(E_{\text{cut}})$  and  $\log_{10}(\text{DM})$  are both particularly discrepant with the data. These logarithmic correlations are associated with the flux threshold which implies that, the more distant the MSP, the larger the intrinsic luminosity it is required to have in order to have a significant probability of being resolved. The point distinguishing Model A1 from Model A9 in regards to the relationship between distance and  $E_{\text{cut}}$  arises because Model A1 has a significant positive  $a_\gamma = 1.2 \pm 0.3$  in Eq. 2.24, while in Model A9 we have set  $a_\gamma = 0$ . Regarding the logarithmic correlation between distance and  $P$ , it follows from the presence of  $P$  in the denominator of Eq. 2.17 that the significantly positive  $d_\gamma = 0.5 \pm 0.1$  in Eq. 2.24 leads to the negative correlation with the distance measures in Model A1, while, again, in Model A9 we have  $d_\gamma = 0$  by construction.

Model A6 is related to model A1 as follows from Eqs. 2.17 and 2.18 which demonstrate that we can go from the luminosity function in Model A1 to the luminosity function in Model A6 by setting  $b_\gamma = \frac{\alpha_\gamma}{2} + \frac{3\beta_\gamma}{2}$  and  $d_\gamma = \frac{\beta_\gamma}{4} - \frac{\alpha_\gamma}{4}$ . However, Model A6 does not have a dependence on  $E_{\text{cut}}$  and therefore assumes  $a_\gamma = 0$ . As can be seen in Table 2.3,  $a_\gamma = 1.2 \pm 0.3$  is significantly positive for Model A1 and so Model A1 is preferred over Model A6. This is consistent with Model A6 having a  $\Delta\text{WAIC} = 11.0$  relative to Model A1. As can be seen in Table 2.5 the problem with Model A6 is that it does not predict the observed logarithmic correlation between distance and  $E_{\text{cut}}$ .

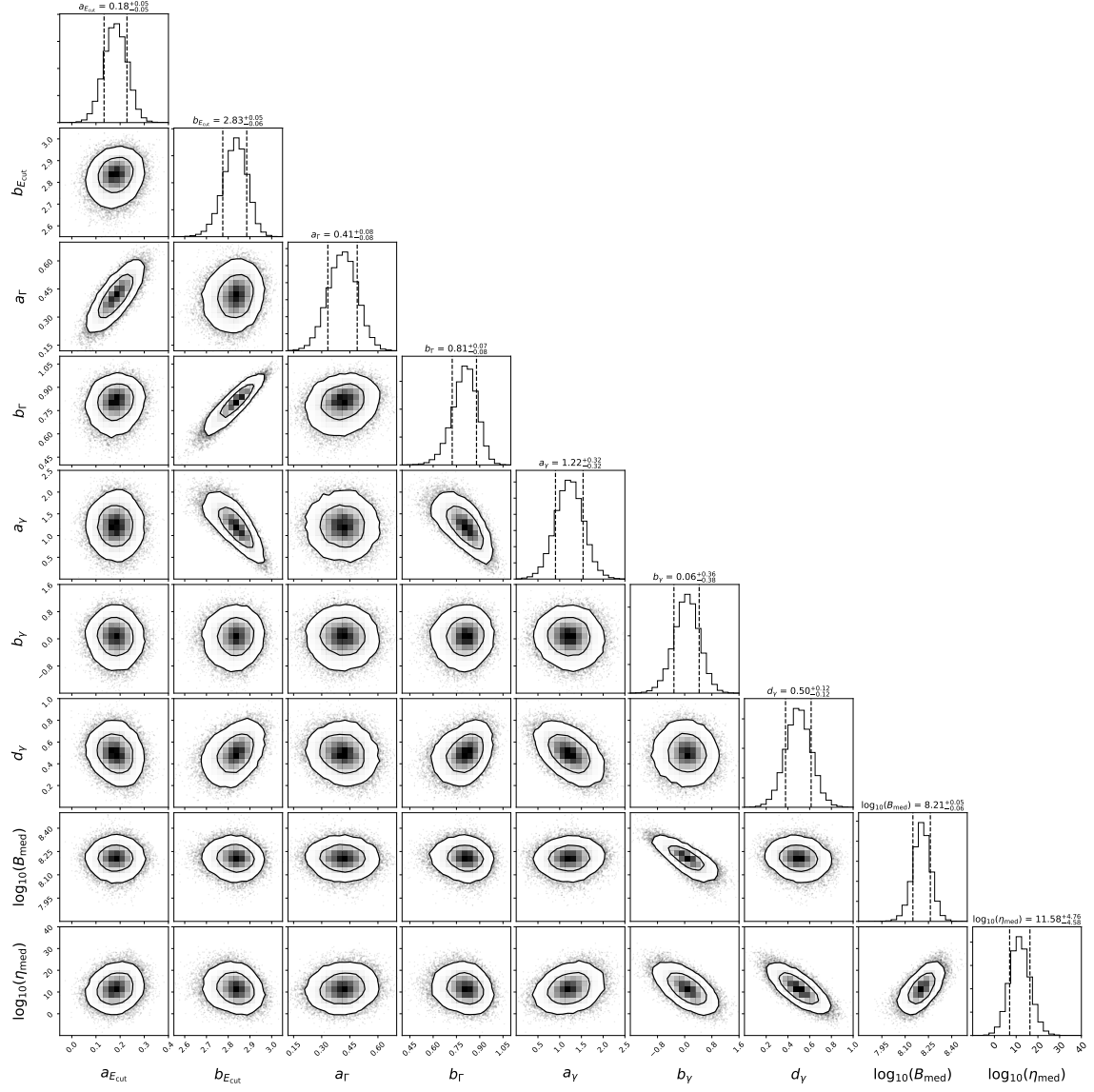


Fig. 2.2 Corner plot showing a selection of parameters for Model A1 ( $L = \eta E_{\text{cut}}^{a_{\gamma}} B^{b_{\gamma}} \dot{E}^{d_{\gamma}}$ ) with 68% and 95% contours. These parameters relate to the luminosity,  $E_{\text{cut}}$ ,  $\Gamma$  and  $B$  distributions. The  $B_{\text{med}}$  parameter has units of Gauss.

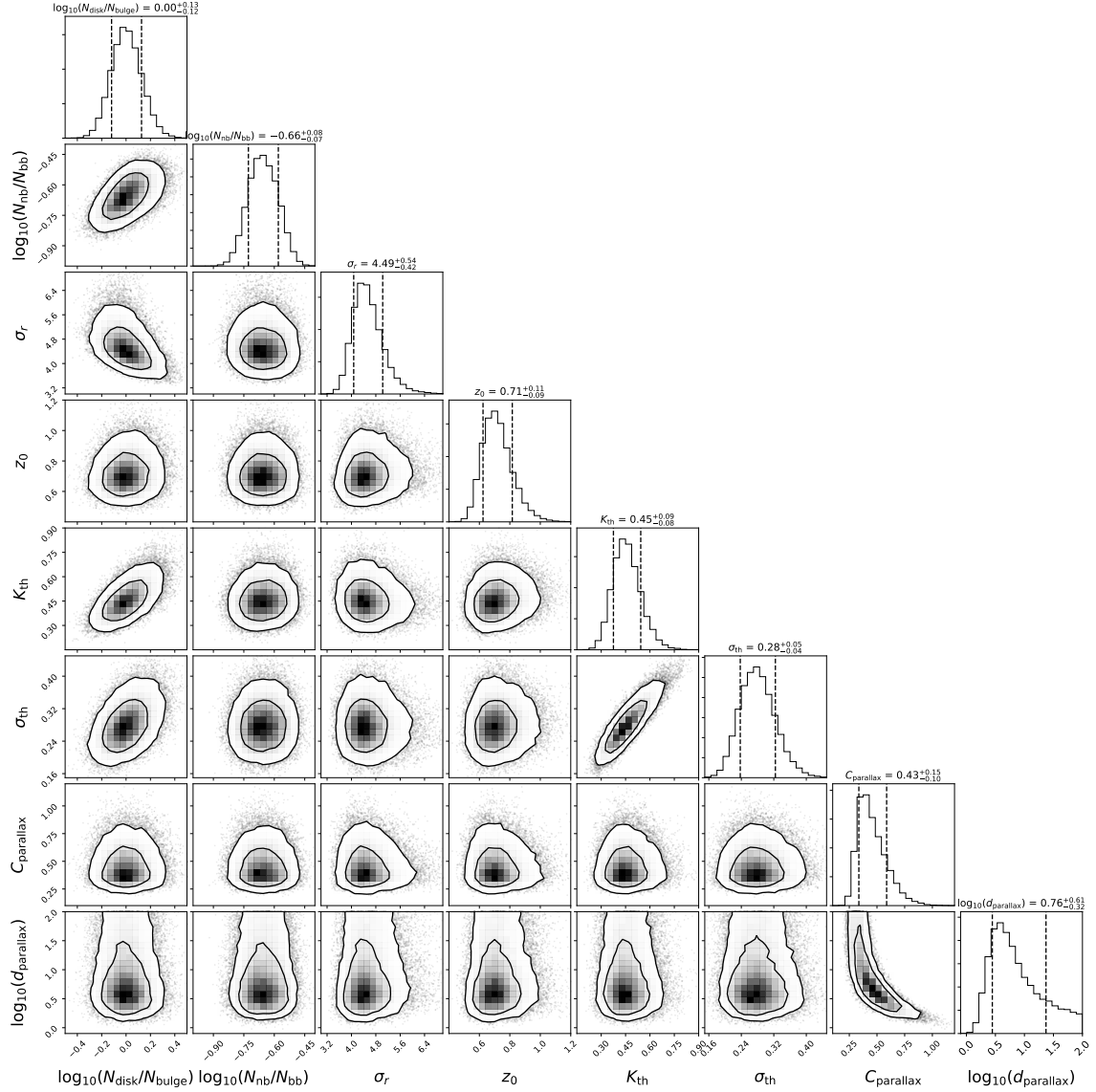


Fig. 2.3 Corner plot showing a selection of parameters for Model A1 with 68% and 95% contours. These parameters relate to the number of MSPs in different components of the spatial distribution, the flux threshold, the model of the parallax measurement availability and the initial period distribution. The parameters  $\sigma_r$ ,  $z_0$ , and  $d_{\text{parallax}}$  have units of kpc.

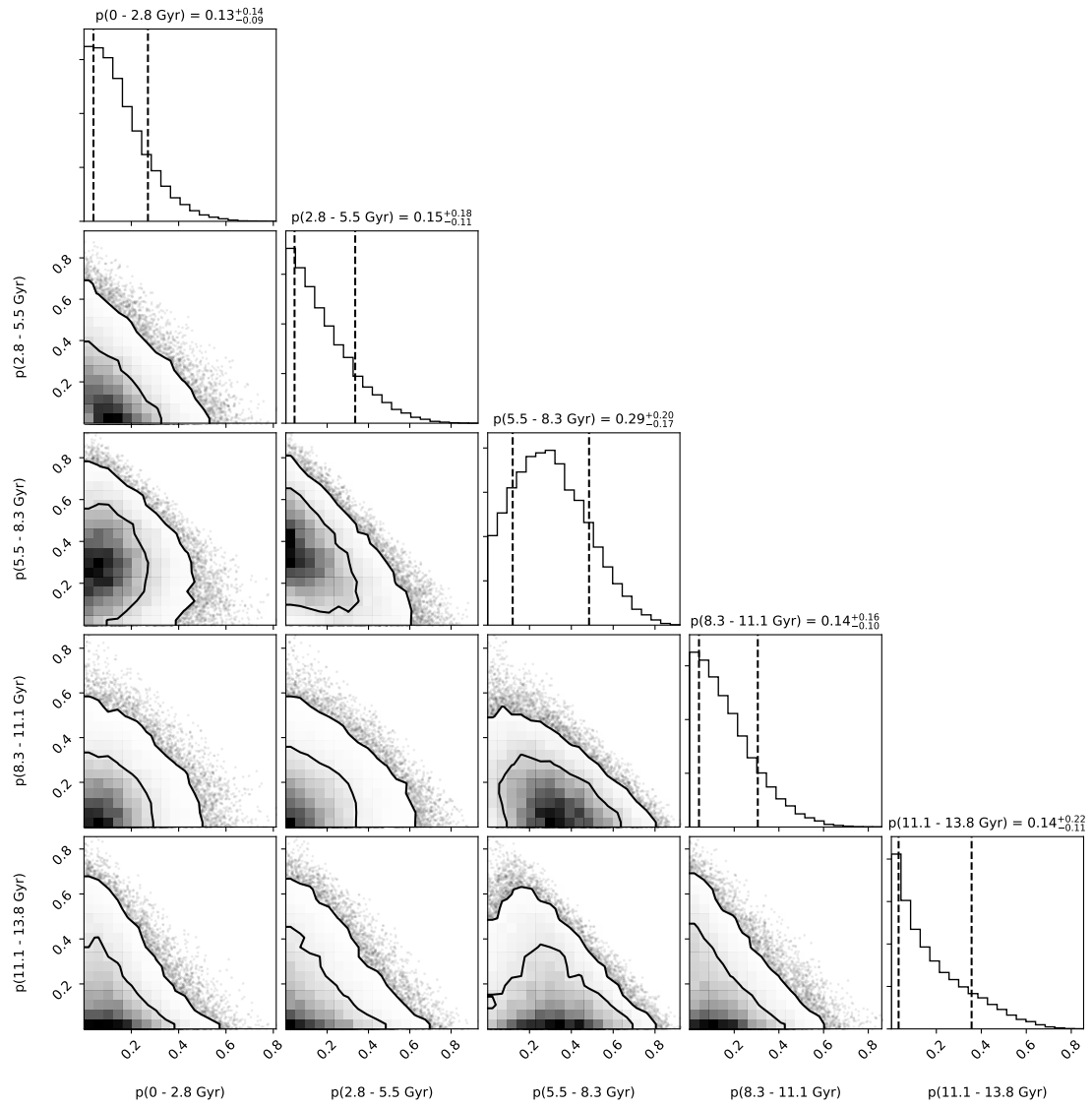


Fig. 2.4 Model A1 probability of delay time within each DTD bin.

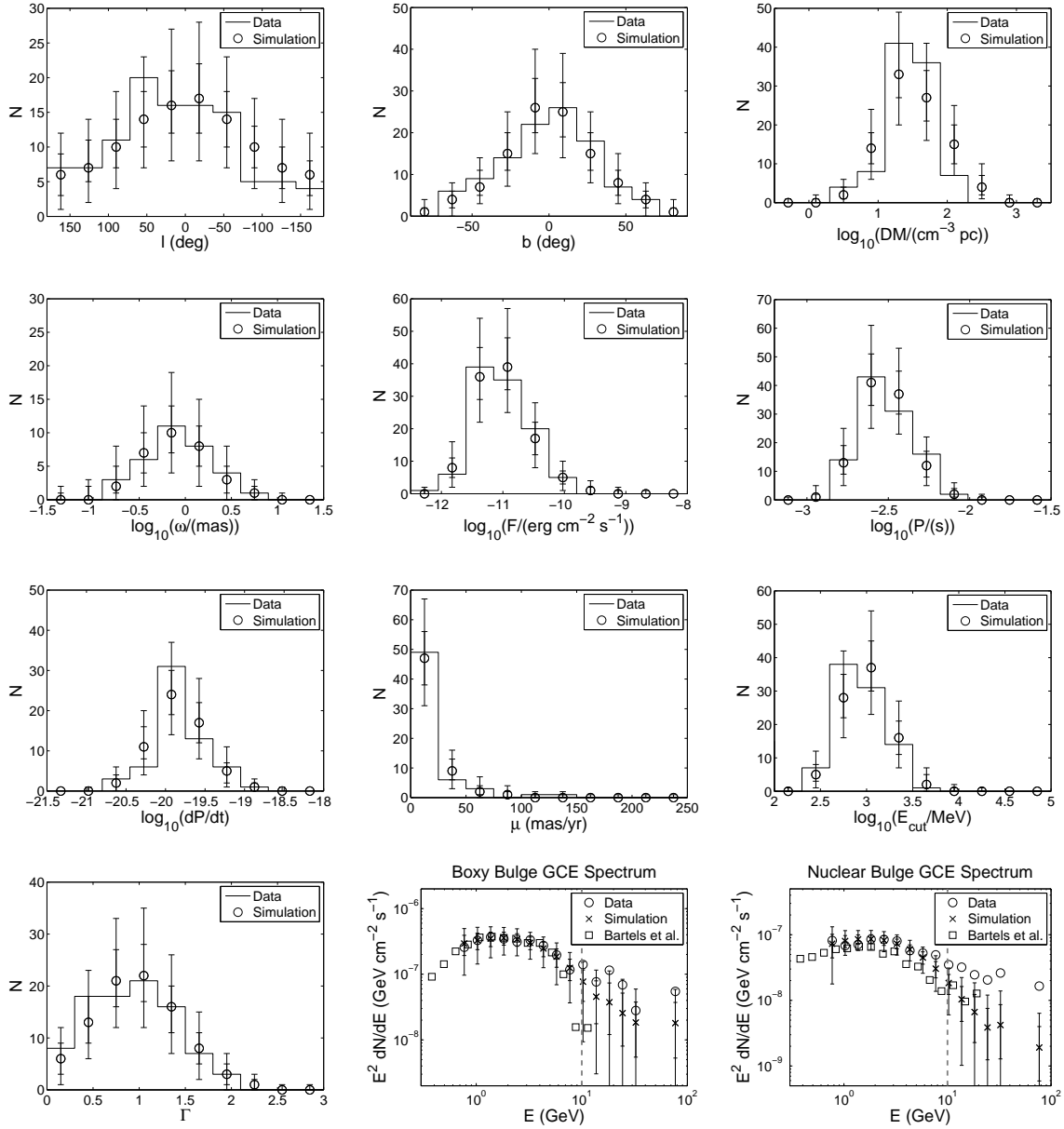


Fig. 2.5 Observed data compared to simulated observations for Model A1 of longitude  $l$  (deg), latitude  $b$  (deg), dispersion measure  $DM$  ( $\text{cm}^{-3} \text{ pc}$ ), parallax  $\omega$  (mas), flux  $F$  ( $\text{erg cm}^{-2} \text{ s}^{-1}$ ), period  $P$  (s), period derivative  $\dot{P}$ , proper motion  $\mu$  ( $\text{mas yr}^{-1}$ ), spectral cutoff  $E_{\text{cut}}$  (MeV), spectral index  $\Gamma$  and GCE spectra for the boxy and nuclear bulges. The GCE spectra are for the inner  $40^\circ \times 40^\circ$  region and the vertical dashed line at 10 GeV shows the maximum energy up to which we fitted. The simulated data are shown as medians, 68% and 95% intervals in each bin. The intervals on the simulated GCE data include the errors on the observed data. For the GCE plots, we also show boxy bulge and nuclear bulge spectra from Bartels et al. [7] in addition to those we fitted to of Macias et al. [8].

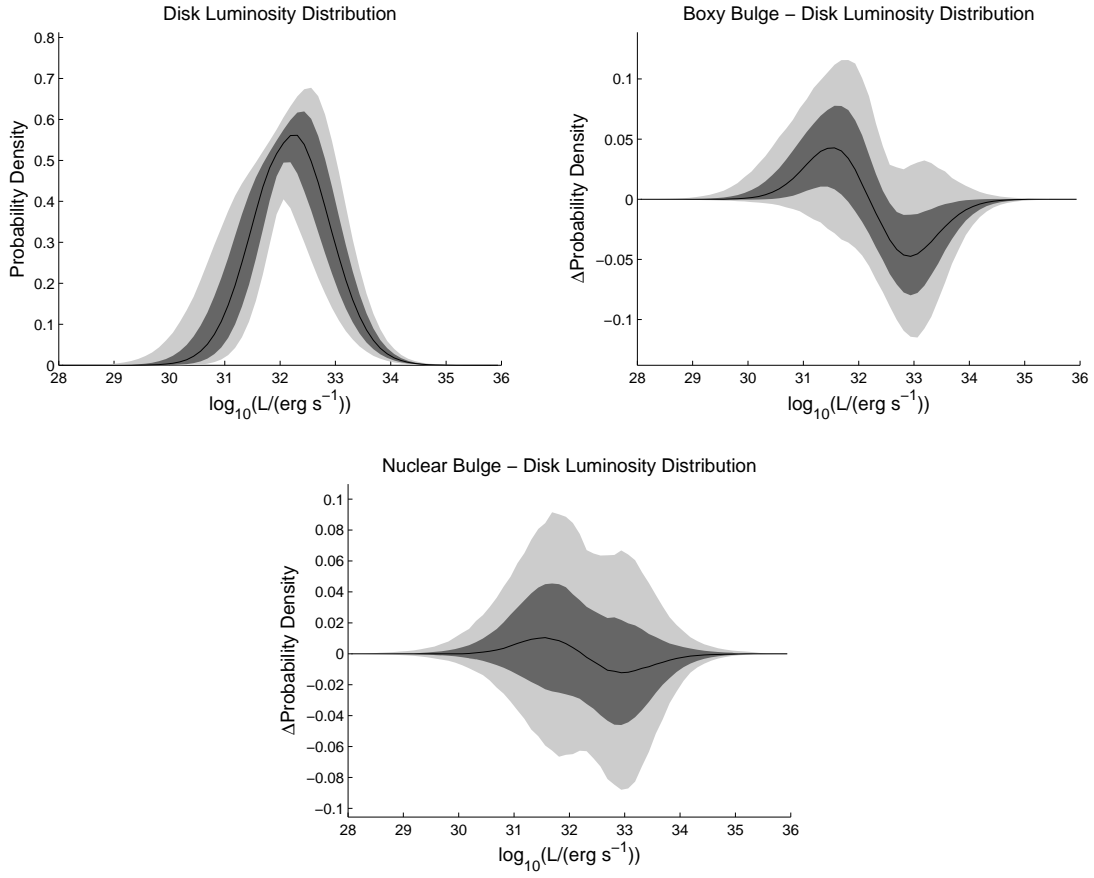


Fig. 2.6 MSP luminosity distribution in the disk and the change in probability density for the boxy bulge and nuclear bulge for Model A1. The black line shows the median probability density in each bin, dark grey the 68% interval and light grey the 95% interval. Note that the probability density is of  $\log_{10}(L/(\text{erg s}^{-1}))$  rather than of  $L/(\text{erg s}^{-1})$ .

$\log_{10}(\omega)$	0.417						
$\log_{10}(F)$	0.627	0.159					
$\log_{10}(P)$	0.126	0.948	0.752				
$\log_{10}(\dot{P})$	0.911	0.245	0.451	0.159			
$\mu$	0.034	0.398	0.294	0.938	0.338		
$\log_{10}(E_{\text{cut}})$	0.789	0.039	0.437	0.527	0.382	0.477	
$\Gamma$	0.550	0.163	0.773	0.583	0.352	0.707	0.612
	$\log_{10}(\text{DM})$	$\log_{10}(\omega)$	$\log_{10}(F)$	$\log_{10}(P)$	$\log_{10}(\dot{P})$	$\mu$	$\log_{10}(E_{\text{cut}})$

Table 2.4 Posterior predictive p-values for the correlation coefficients between observables for Model A1 which has  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ .

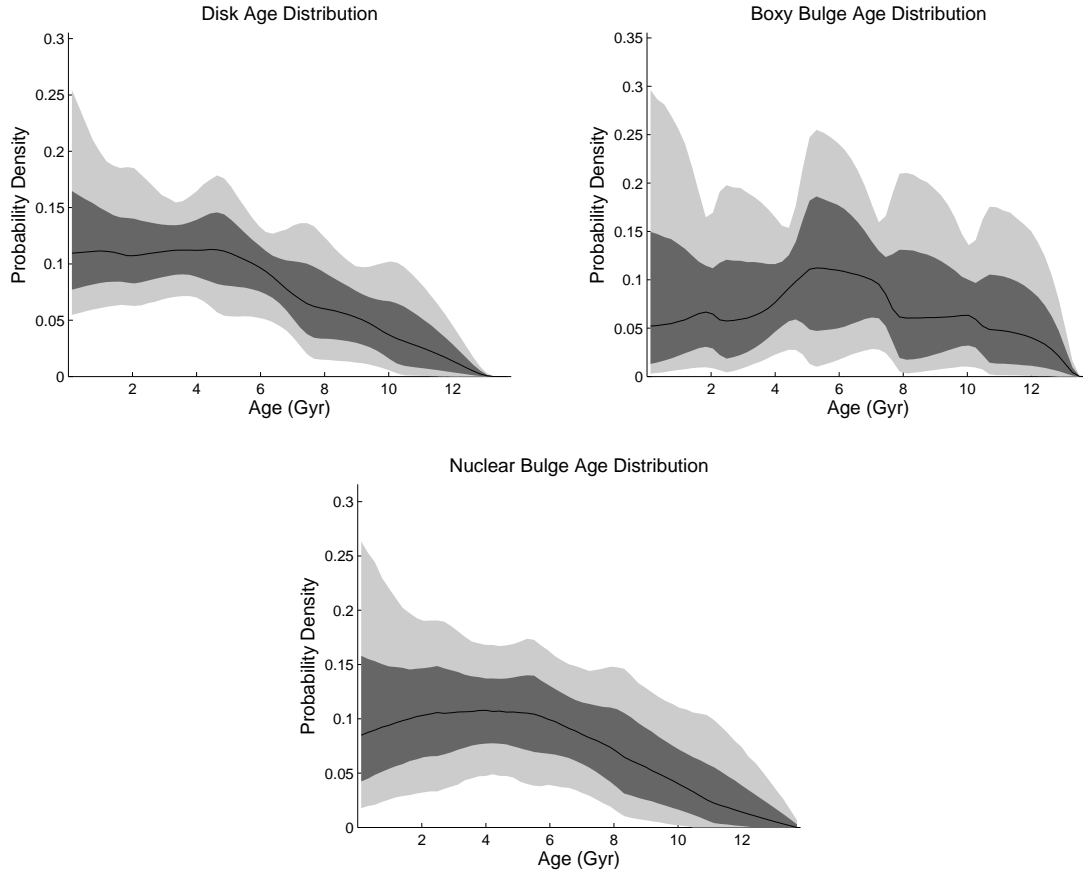


Fig. 2.7 MSP age distribution in the disk, boxy bulge and nuclear bulge for Model A1. The black line shows the median probability density in each bin, dark grey the 68% interval and light grey the 95% interval.

$\log_{10}(\omega)$	0.426						
$\log_{10}(F)$	0.615	0.158					
$\log_{10}(P)$	0.138	0.941	0.743				
$\log_{10}(\dot{P})$	0.891	0.265	0.431	0.145			
$\mu$	0.430	0.395	0.282	0.932	0.366		
$\log_{10}(E_{\text{cut}})$	0.992	0.007	0.697	0.609	0.216	0.175	
$\Gamma$	0.852	0.069	0.879	0.667	0.223	0.515	0.625
	$\log_{10}(\text{DM})$	$\log_{10}(\omega)$	$\log_{10}(F)$	$\log_{10}(P)$	$\log_{10}(\dot{P})$	$\mu$	$\log_{10}(E_{\text{cut}})$

Table 2.5 Posterior predictive p-values for the correlation coefficients between observables for Model A6 which has  $L = \eta P^{\alpha_\gamma} \dot{P}^{\beta_\gamma}$ . Values greater than 0.99 and less than 0.01 are marked in red.



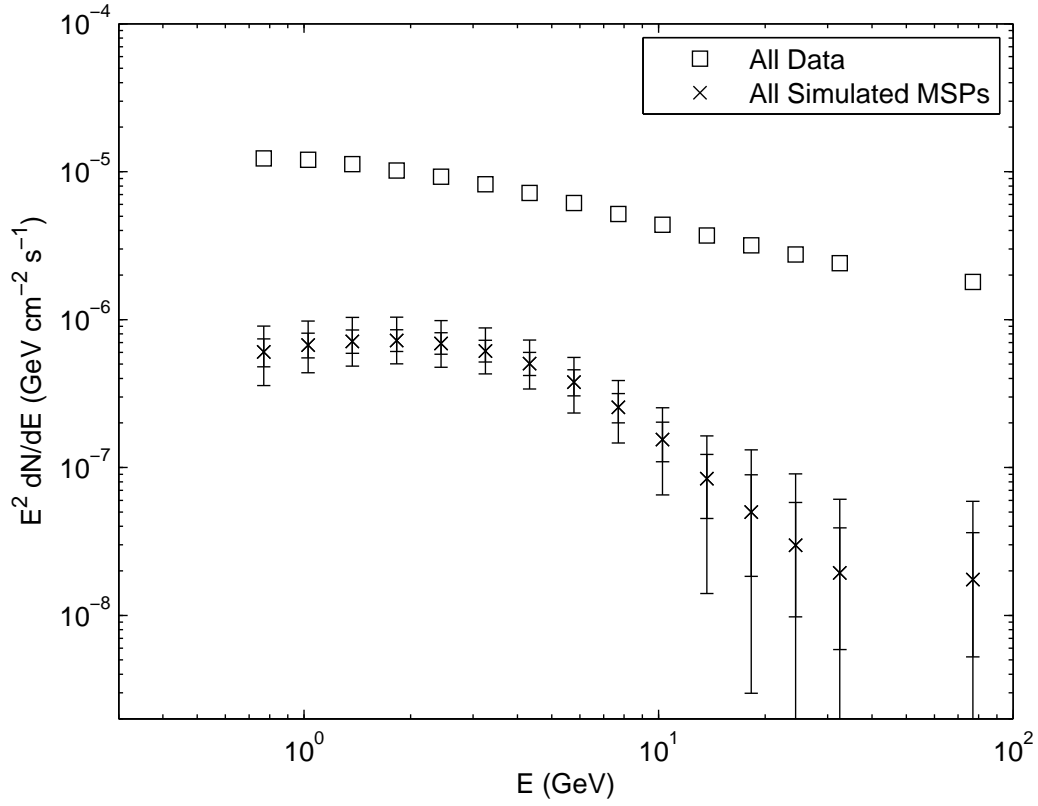


Fig. 2.8 Total observed gamma ray emission compared to all simulated MSP emission from inner  $40^\circ \times 40^\circ$  region for Model A1. Data are from Abazajian et al. [9].

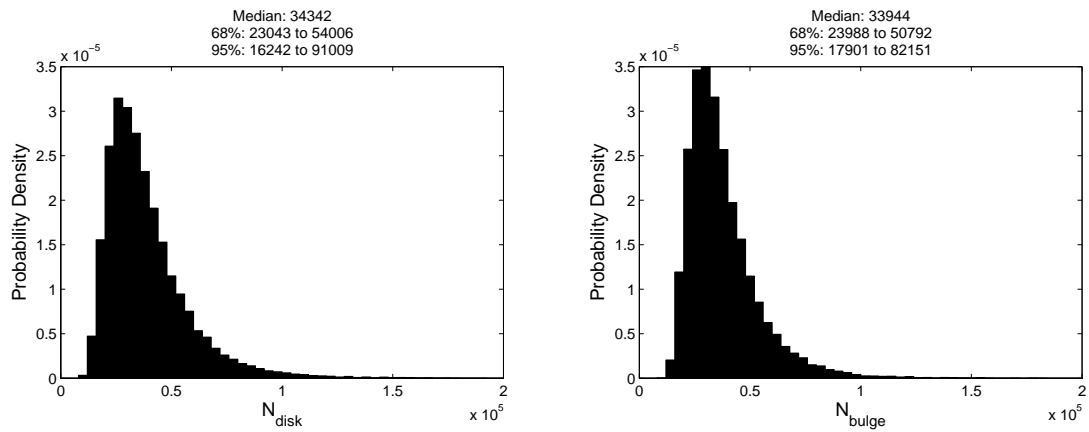


Fig. 2.9 Distribution of the number of disk and bulge MSPs for Model A1.

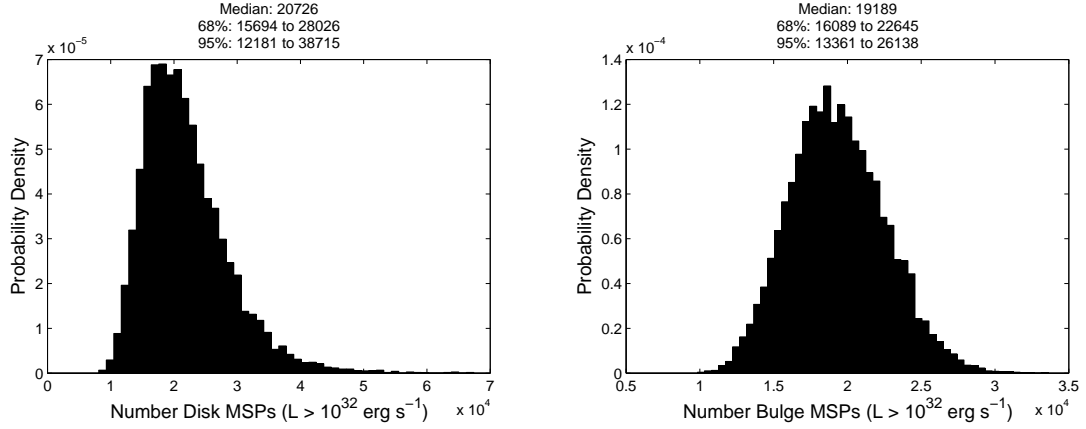


Fig. 2.10 Distribution of the number of disk and bulge MSPs with  $L > 10^{32} \text{ erg s}^{-1}$  for Model A1.

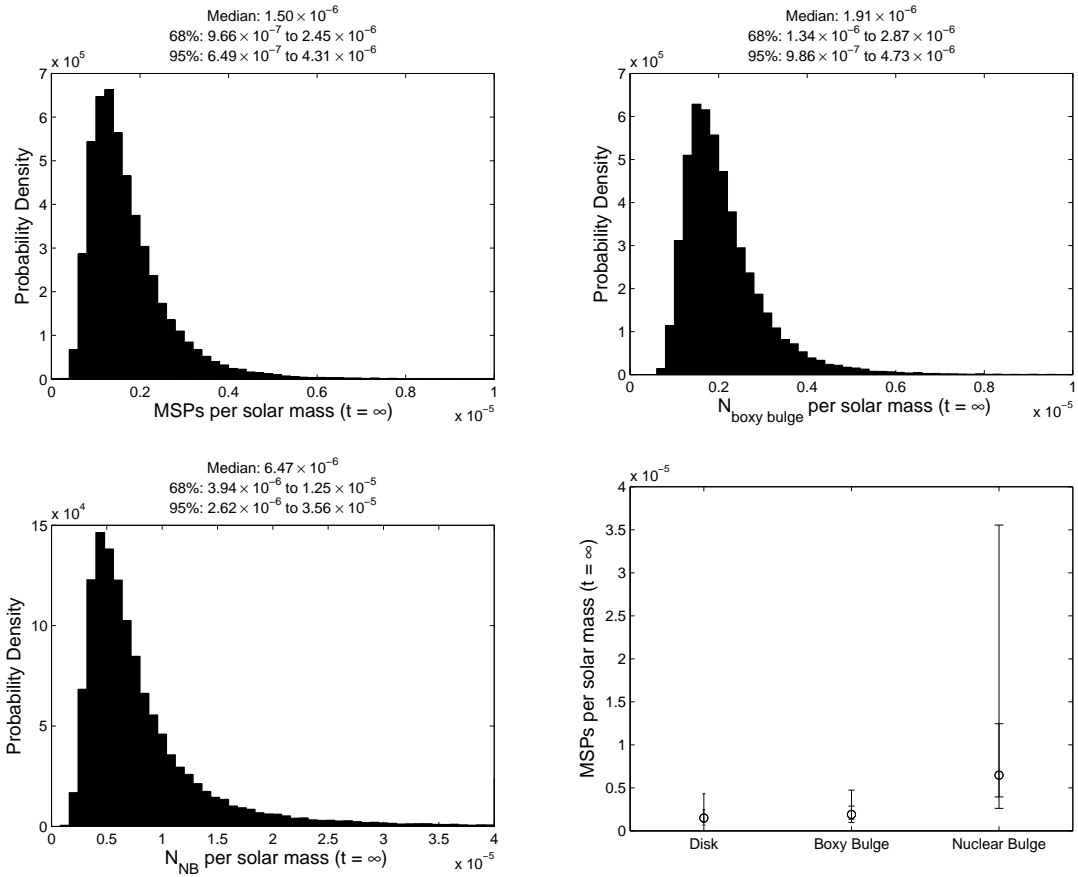


Fig. 2.11 Distribution of the number of disk, boxy bulge and nuclear bulge MSPs produced per solar mass at  $t = \infty$  assuming no further star formation beyond today for Model A1. In the bottom right plot, the medians, 68% and 95% intervals are shown side by side.

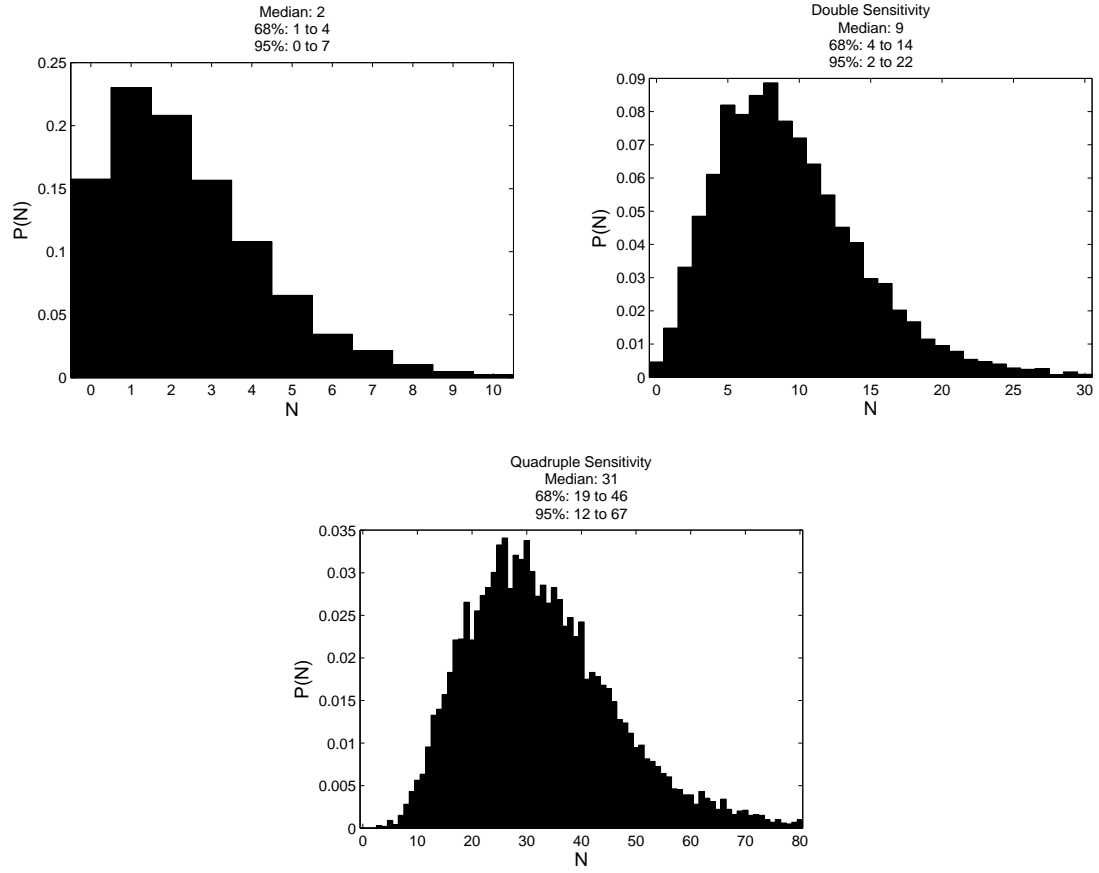


Fig. 2.12 Probability of observing  $N$  bulge MSPs with the current sensitivity as well as for two and four times the current sensitivity for Model A1. The double and quadruple sensitivity distributions were produced by subtracting 0.3 and 0.6 respectively from  $K_{\text{th}}$ .

$\log_{10}(\omega)$	0.384						
$\log_{10}(F)$	0.719	0.120					
$\log_{10}(P)$	0.057	0.965	0.646				
$\log_{10}(\dot{P})$	0.938	0.225	0.502	0.157			
$\mu$	0.412	0.415	0.244	0.963	0.316		
$\log_{10}(E_{\text{cut}})$	0.942	0.017	0.583	0.004	0.819	0.339	
$\Gamma$	0.889	0.059	0.906	0.000	0.894	0.509	0.661
	$\log_{10}(\text{DM})$	$\log_{10}(\omega)$	$\log_{10}(F)$	$\log_{10}(P)$	$\log_{10}(\dot{P})$	$\mu$	$\log_{10}(E_{\text{cut}})$

Table 2.6 Same as Table 2.5 except for Model A7 which has  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ ,  $a_{E_{\text{cut}}} = a_\Gamma = 0$ .

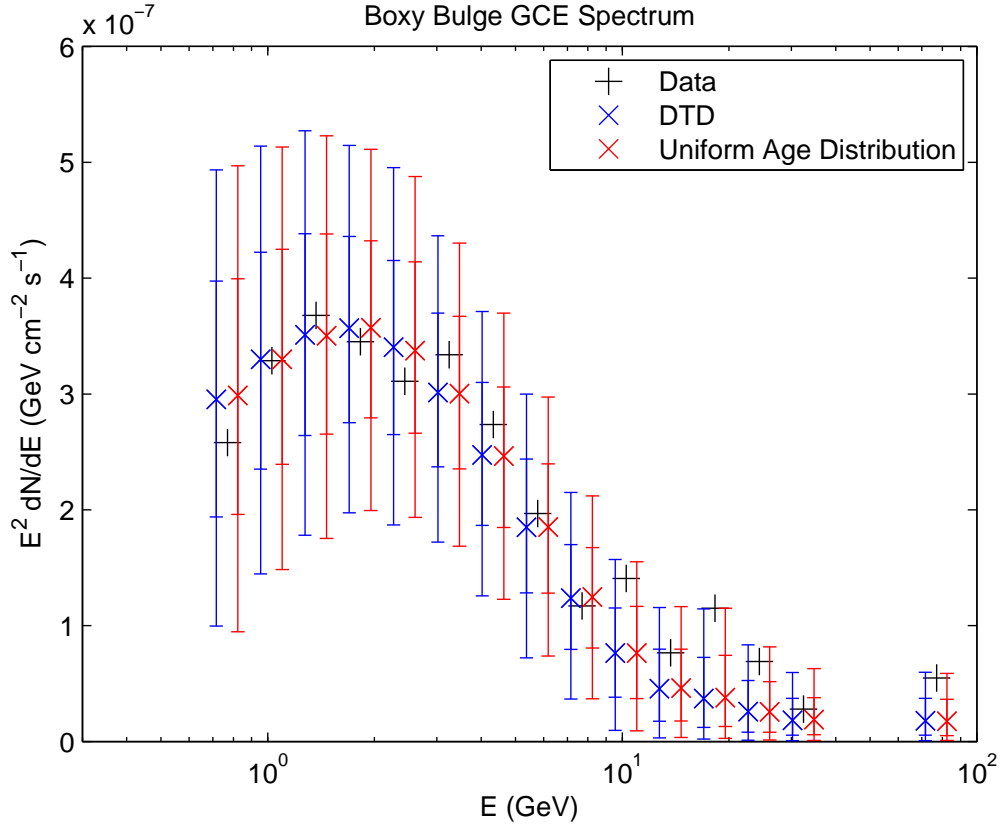


Fig. 2.13 Comparison of posterior predictive boxy bulge GCE spectra for the  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$  model in the case where we have a DTD (Model A1) and the case where a uniform MSP age distribution was used (Model A2). The DTD case has been shifted slightly to the left and the uniform case to the right in order to make comparison easier. Only bins with  $E < 10$  GeV were fitted.

$\log_{10}(\omega)$	0.422						
$\log_{10}(F)$	0.670	0.141					
$\log_{10}(P)$	<b>0.000</b>	<b>0.998</b>	0.289				
$\log_{10}(\dot{P})$	<b>0.997</b>	0.104	0.697	0.140			
$\mu$	0.424	0.396	0.275	<b>0.998</b>	0.155		
$\log_{10}(E_{\text{cut}})$	<b>1.000</b>	<b>0.002</b>	0.869	0.541	0.316	0.053	
$\Gamma$	<b>0.998</b>	0.012	0.978	0.621	0.354	0.192	0.631
	$\log_{10}(\text{DM})$	$\log_{10}(\omega)$	$\log_{10}(F)$	$\log_{10}(P)$	$\log_{10}(\dot{P})$	$\mu$	$\log_{10}(E_{\text{cut}})$

Table 2.7 Same as Table 2.5 except for Model A9 which has  $L = \eta$ .

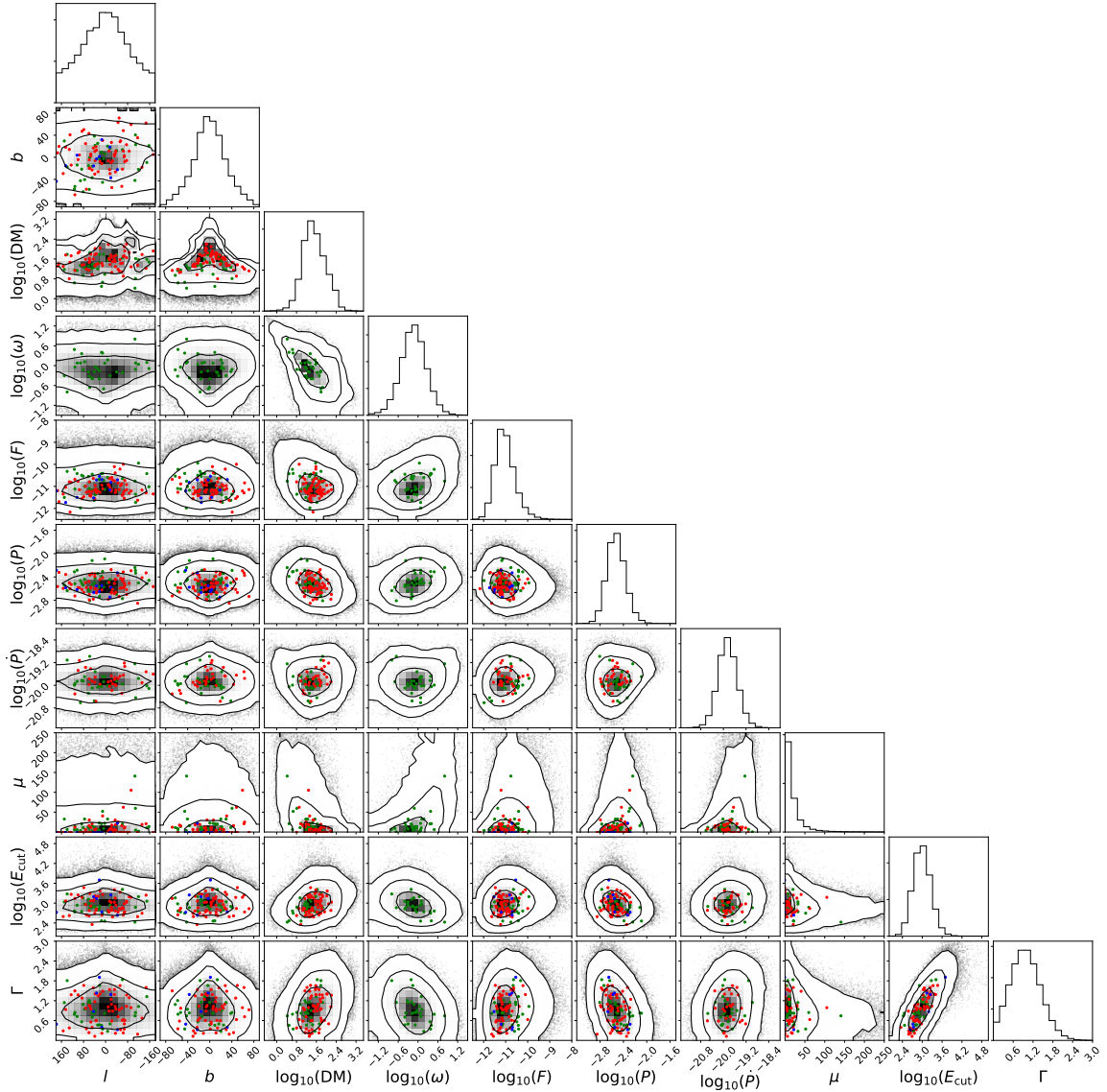


Fig. 2.14 Model A1 ( $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ ) posterior predictive corner plots showing longitude  $l$  (deg), latitude  $b$  (deg), dispersion measure DM ( $\text{cm}^{-3}$  pc), parallax  $\omega$  (mas), flux  $F$  ( $\text{erg cm}^{-2} \text{s}^{-1}$ ), period  $P$  (s), period derivative  $\dot{P}$ , proper motion  $\mu$  ( $\text{mas yr}^{-1}$ ), spectral cutoff  $E_{\text{cut}}$  (MeV) and spectral index  $\Gamma$  simulated and real data. Red points are MSPs with dispersion measure distances, green is parallax, and blue is no distance measurement. The contours show regions containing 68%, 95% and 99.7% of the simulated data.

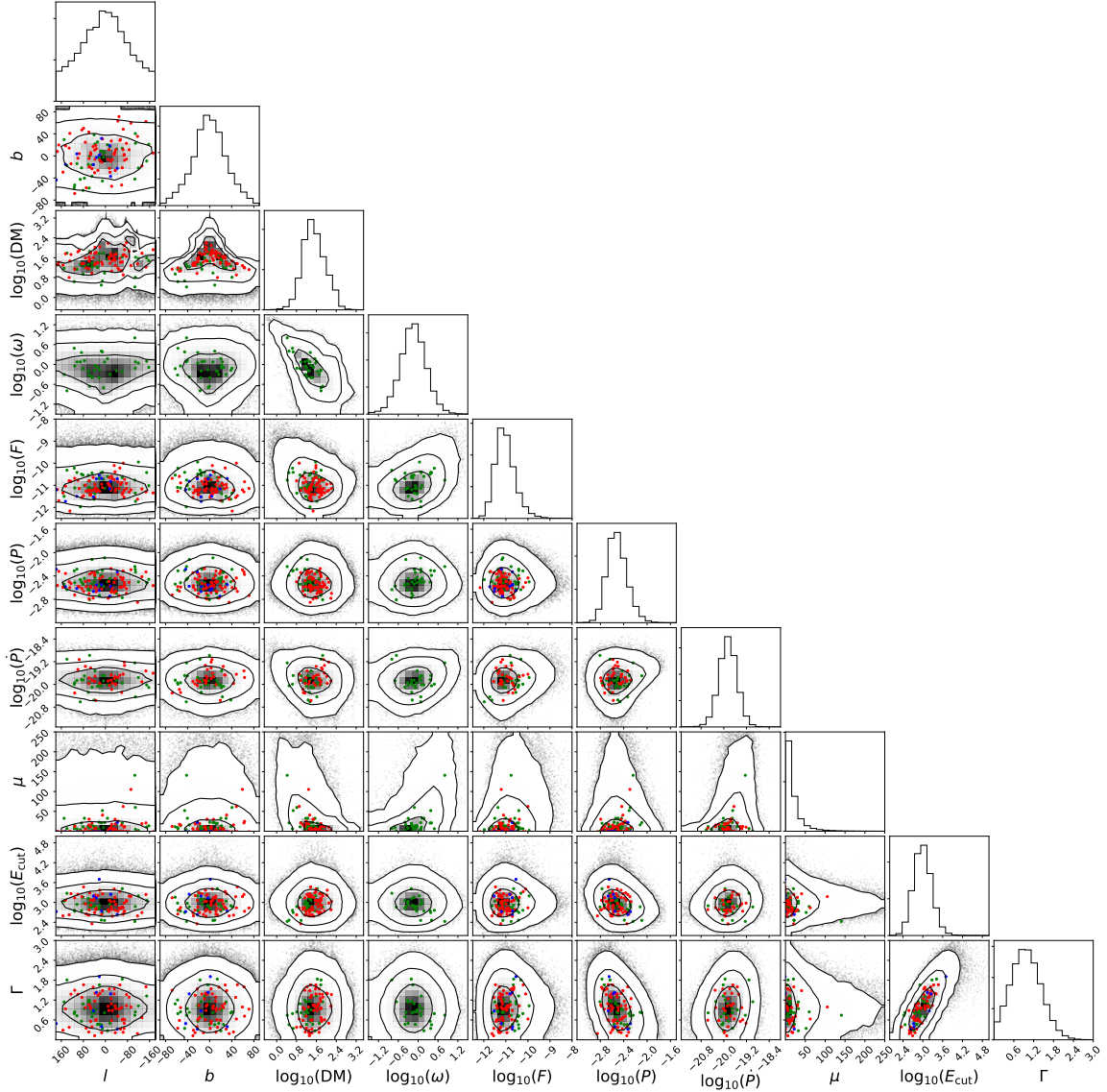


Fig. 2.15 The same as Fig. 2.14 except for Model A9 which has  $L = \eta$ .

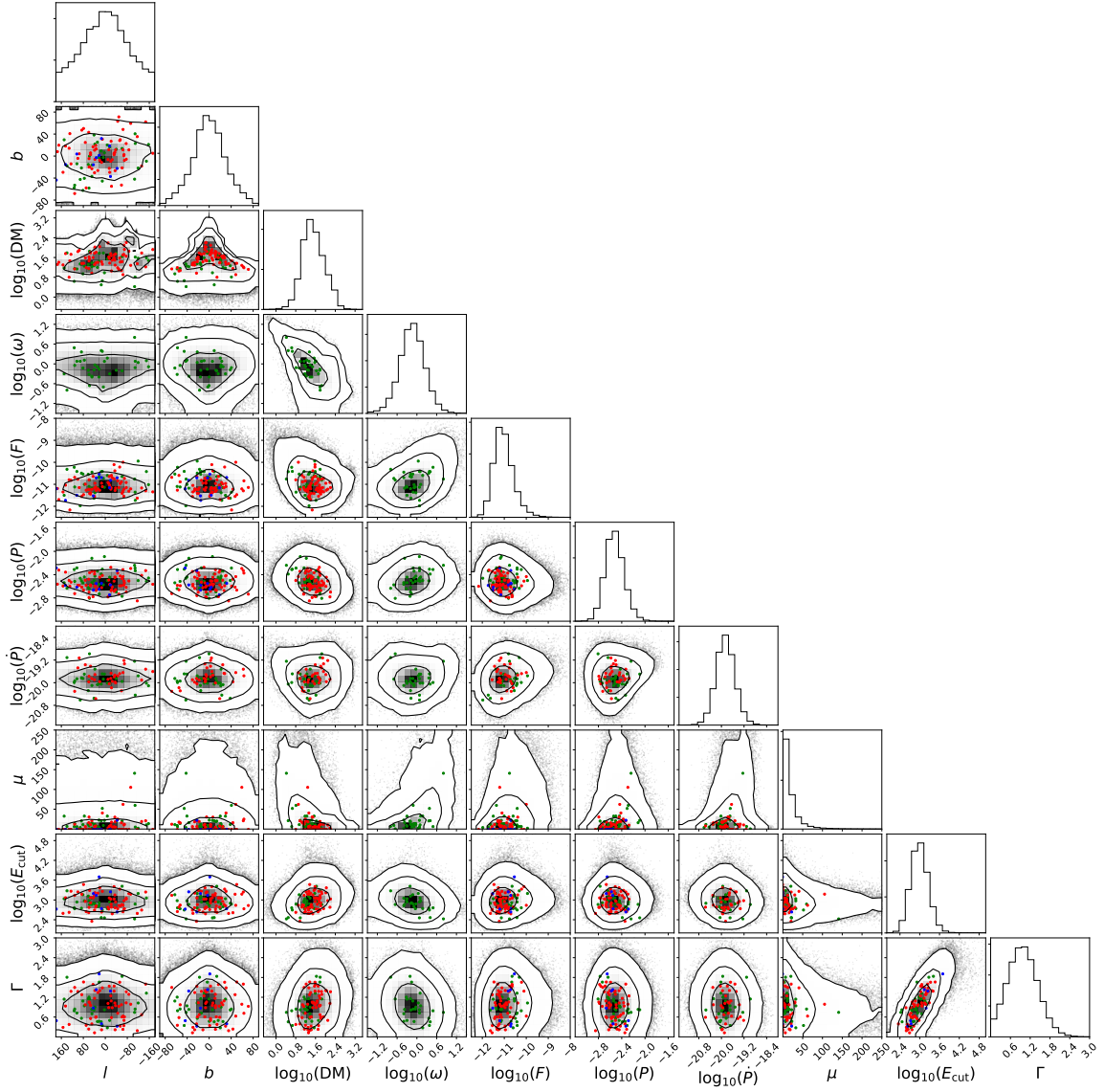


Fig. 2.16 The same as Fig. 2.14 except for Model A7 which has the same luminosity function as Model A1 but has  $E_{\text{cut}}$  independent of  $\dot{E}$  ( $a_{E_{\text{cut}}} = a_{\Gamma} = 0$ ).

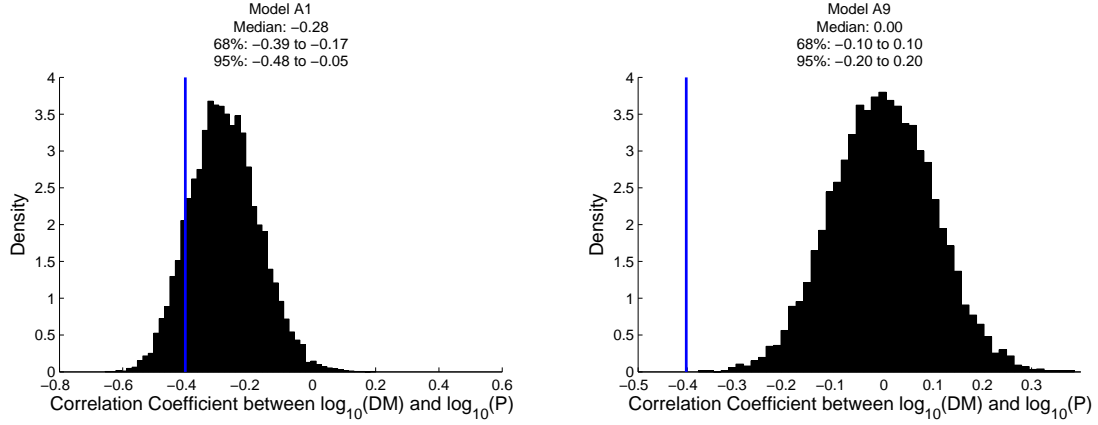


Fig. 2.17 Comparison of correlation coefficients between  $\log_{10}(P)$  and  $\log_{10}(\text{DM})$  for resolved MSP simulated and real data for Model A1 ( $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ ) on the left and Model A9 ( $L = \eta$ ) on the right. The blue line shows the correlation coefficient for the resolved MSP data at -0.40. The posterior predictive p-values are 0.13 for A1, and 0 for A9.

Model A7 is the same as Model A1 except that we set  $a_{E_{\text{cut}}} = a_\Gamma = 0$ . We see from Table 2.3 that this choice for Model A7 is disfavoured as we found that Model A1 had a significantly positive  $a_\Gamma = 0.41 \pm 0.08$  for Eq. 2.29. This preference for Model A1 over A7 is confirmed in Table 2.2 where it can be seen that Model A7 has a  $\Delta\text{WAIC} = 19.4$  relative to Model A1. We can see how Model A7 produced a worse fit to the data by comparing Figs. 2.14 and 2.16: in the latter case where  $a_{E_{\text{cut}}} = a_\Gamma = 0$ , the relationship between spectral index and period has clearly disappeared. This is confirmed in Table 2.6 where two of the correlation coefficients are of concern and, in particular, the correlation coefficient between  $\Gamma$  and  $\log_{10}(P)$  has a posterior predictive p-value of 0.000.

Our best fitting model of the Galactic MSP population was Model A1 for which  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ . From inspection of Figs. 2.5 and 2.14 it is evident that this model generally provided a good fit to the resolved MSP data. This is also confirmed in Table 2.4 where all the posterior predictive p-values are within the 1% to 99% range. We find  $a_\gamma = 1.2^{+0.3}_{-0.3}$ ,  $b_\gamma = 0.1^{+0.4}_{-0.4}$  and  $d_\gamma = 0.5^{+0.1}_{-0.1}$ . These results are consistent with those obtained by Kalapotharakos et al. [18]. These authors performed a least squares fit to both the MSPs and young pulsars in the Second Fermi Pulsar Catalog [45] and find  $a_\gamma = 1.12 \pm 0.24$ ,  $b_\gamma = 0.17 \pm 0.05$  and  $d_\gamma = 0.41 \pm 0.08$ . Kalapotharakos et al. [18] point out that their results are consistent with predicted values of  $a_\gamma = 4/3$ ,  $b_\gamma = 1/6$  and  $d_\gamma = 5/12$  in the case that curvature radiation is the source of gamma-ray emission. This stands in contrast to the case of synchrotron radiation for which  $a_\gamma = 1$ ,  $b_\gamma = 0$  and  $d_\gamma = 1$  is expected. Our posterior distributions for  $a_\gamma$  and  $b_\gamma$  are consistent with both cases to within  $2\sigma$ , but our  $d_\gamma$  is inconsistent with synchrotron



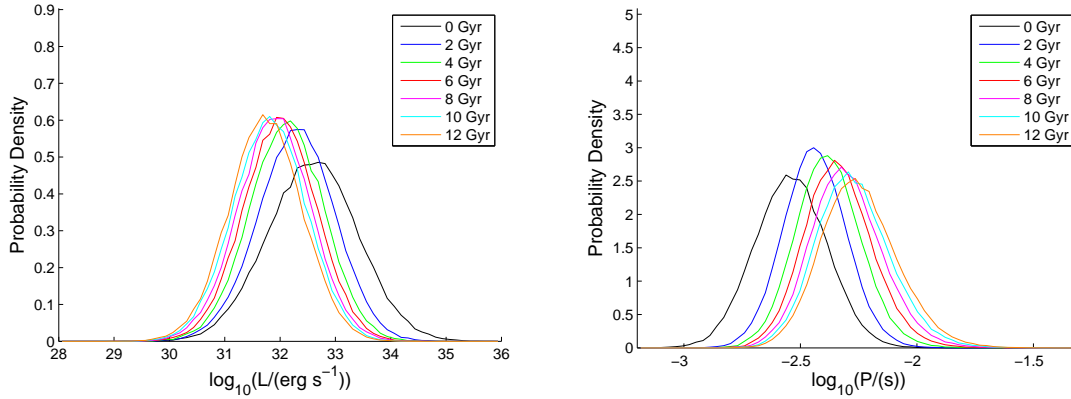


Fig. 2.18 Evolution of the luminosity and period distributions with time for Model A1 which has  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ . The legend gives the age of the MSPs in the corresponding distribution. These figures were made by selecting randomly from the highest likelihood parameter sets in our eight Markov chains for this model.

radiation. One major difference between the work of Kalapotharakos et al. [18] and ours is that we use many more MSPs, but no young pulsars. Overall we have a similar total number of pulsars. However, because of the Shklovskii effect, the intrinsic period derivative  $\dot{P}_{\text{int}}$  is poorly determined relative to the case of young pulsars. Also,  $B$  and  $\dot{E}$  depend on  $\dot{P}_{\text{int}}$  through Eqs. 2.18 and 2.17 respectively. It follows that some MSPs with smaller values of  $\dot{P}$  may have poorly-determined period derivatives relative to the case presented by young pulsars.

In Model A1, which has  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , relatively young members of the underlying population of MSPs are more likely to be resolved as they are brighter. This follows as  $\dot{E}$ , and therefore luminosity, decreases with age. As the magnetic field strength is assumed constant for each MSP over time, it can be seen from Eqs. 2.18 and 2.17 that  $\dot{E} \propto P^{-4}$ . We show in Fig. 2.18 an example of how the  $\log_{10}(L)$  and  $\log_{10}(P)$  probability density functions evolve with age.

As can be seen from Fig. 2.4, the DTD peaks in the central 5.5 to 8.3 Gyr bin producing an age distribution that tends to plateau starting around 5 Gyr ago as shown in Fig. 2.7. It can be seen in Table 2.2 that there is no significant difference in the WAIC for the DTD versus the uniform age distribution case. As can be seen in Fig. 2.7, the fitted DTD produces an age distribution that is similar to the uniform case. For the uniform age distribution, the probability density is 0.1 for ages less than 10 Gyr, and 0 elsewhere.

The models that were significantly worse when the GCE was included remained significantly worse when it was not. The model for which the luminosity obeys  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$  remained the best model whether the GCE was included or not, i.e., Models A1 and A2 were

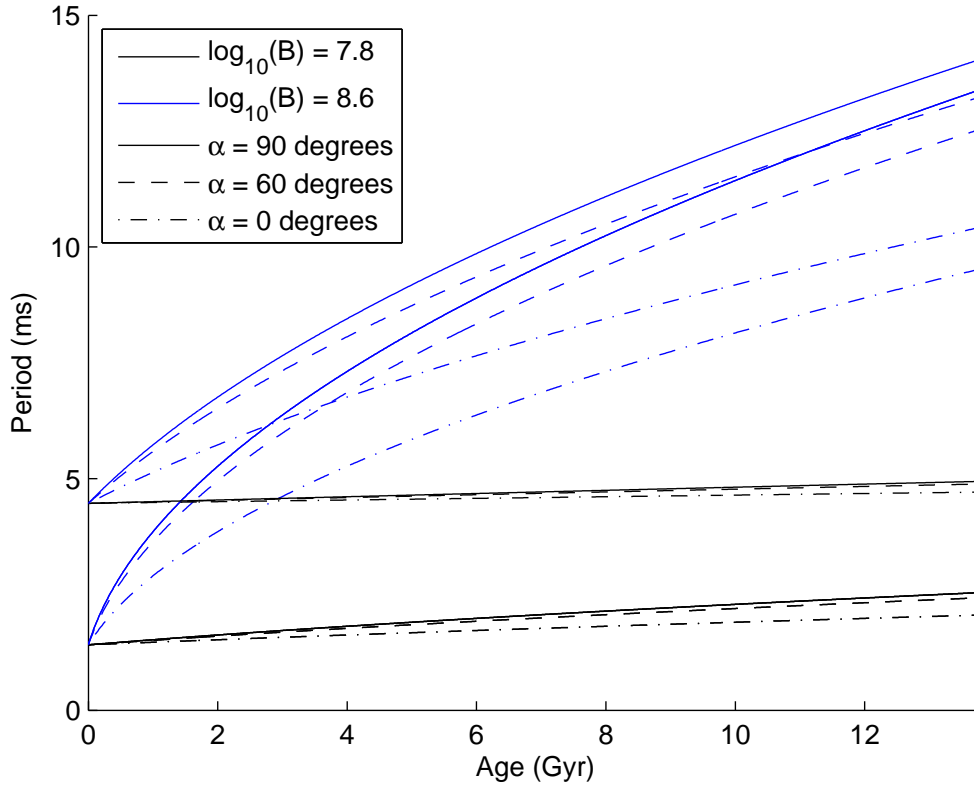


Fig. 2.19 Evolution of period with time for various magnetic field strengths  $B$ , magnetic field axis angles  $\alpha$  and initial periods  $P_i$ . The two  $P_i$  values used were  $10^{-2.85}$  s and  $10^{-2.35}$  s.

the best models when the GCE was included and Models B1 and B2 were the best models when the GCE was not included. As can be seen in Table 2.3 the posterior distributions of the parameters are similar for the GCE and No GCE cases. All the fitted parameters of Model B2 are consistent with those of Model A1, which is the same except the former lacks a bulge MSP population and GCE.

As can be seen from Fig. 2.13, there is a small difference in the GCE spectrum for the Model A1 and A2. This is due to the spectral dependence on  $\dot{E}$  in Eqs. 2.28 and 2.29 and also the different star formation rates for the bulge and disk which are illustrated in Fig. 2.1. However, as can be seen in Fig. 2.13, the differences in the predicted spectrum are negligible in comparison to the model prediction uncertainties. As can be seen in the top right panel of Fig. 2.6 the boxy bulge MSPs do have a median luminosity function that is less bright than the disk MSPs. However, the 95% interval band encompasses zero which is the no difference case. This shows that with the current levels of uncertainty the differences in the properties of the bulge and disk MSPs is not significant.

In Fig. 2.19 we show the evolution of period with time (Eq. 2.23) for MSPs with different properties. This indicates where the constraints on our models of the initial period and age distribution come from. For a resolved MSP with a weak magnetic field, the current period will be near the initial period even if the MSP is old, so these MSPs should be approximately distributed like the initial period distribution. This is a consequence of the fact that an MSP cannot be older than the universe. On the other hand, MSPs with strong magnetic fields will quickly, within a couple of Gyr, move out of the initial period distribution. This (relatively) rapid evolution of period means the distributions of  $P_I$  and  $\alpha$  produce (for a given  $P$  and  $B$ ) a range of possible ages for each MSP.

In Fig. 2.9 it is shown that, at 68% confidence level, for Model A1 between around 23000 and 54000 MSPs are needed in the bulge to produce the observed GCE. This is consistent with Gonthier et al. [19] in which they find, though with a different bulge density model and GCE spectrum, 34,200 MSPs are needed with 11,500 in the region of interest associated with the Gordon et al. [5] GCE spectrum. Calore et al. [109], assuming the ratio of gamma-ray flux to number of radio-bright MSPs in globular clusters is the same as that of the bulge, estimate  $(2.7 \pm 0.2) \times 10^3$  radio-bright Galactic bulge MSPs, defined as those with a flux density at 1.4 GHz of  $\geq 10 \mu\text{Jy}$ , and a total number of  $(9.2 \pm 3.1) \times 10^3$  assuming their adopted radio luminosity function. In Ploeg et al. [15] it was found that around  $(4.0 \pm 0.9) \times 10^4$  MSPs with  $L > 10^{32} \text{ erg s}^{-1}$  were needed to produce the GCE; here, as shown in Fig. 2.10, the required number is between about 13000 and 26000 at 95% confidence interval. Note the difference between these two estimates is not statistically significant; the slight discrepancy is likely due to the more accurate bulge geometric model, luminosity function, and GCE spectrum used in the current study.

A number of studies have fitted luminosity functions to resolved MSPs. For our fit of an independent log-normal luminosity distribution, Model A9 where  $L = \eta$ , we found a median of  $\log_{10}(L_{\text{med}}) = 32.17^{+0.23}_{-0.31}$  and a standard deviation in  $\log_{10}(L)$  of  $\sigma_L = 0.72^{+0.08}_{-0.06}$ . Similarly, fitting log-normal luminosity distributions to the data, Bartels et al. [16] found  $\log_{10}(L_{\text{med}}) = 32.44^{+0.19}_{-0.22}$  and  $\sigma_L = 0.68^{+0.07}_{-0.06}$ , and Hooper and Mohlabeng [13] found best fit parameters of  $\log_{10}(L_{\text{med}}) = 32.67$  and  $\sigma_L = 0.61$ . For Model A6, where  $L \propto P^{\alpha_\gamma} \dot{P}^{\beta_\gamma}$ , we found  $\alpha_\gamma = -2.2^{+0.4}_{-0.4}$  and  $\beta_\gamma = 0.79^{+0.20}_{-0.20}$ , consistent with Gonthier et al. [19] who found  $(\alpha_\gamma, \beta_\gamma)$  of  $(-2.3 \pm 0.3, 0.8 \pm 0.2)$ ,  $(-2.1 \pm 0.3, 0.7 \pm 0.2)$  and  $(-2.5 \pm 0.3, 0.9 \pm 0.2)$  for three different models of gamma-ray emission geometry.

In Fig. 2.20 we compare, for Model A1, the photon flux distributions of MSPs to the sub-threshold distributions inferred from non-Poissonian template fits to Fermi-LAT photon count data from the inner Galaxy region [48, 51]. Although there remains significant systematic uncertainty, these methods can be used to reconstruct the source count distribution down

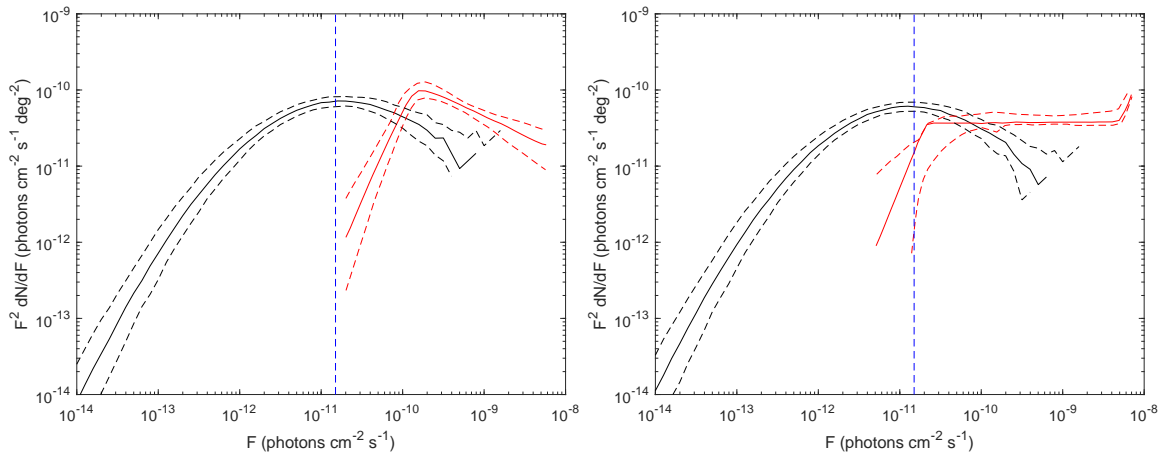


Fig. 2.20 Source count distribution for Model A1 compared to that of two studies applying non-Poissonian template fits to Fermi-LAT data. The reconstructed flux distributions are shown in red, in black is the modelled MSPs, the red and black dashed lines show 68% intervals. The figure on the left is the photon flux distribution compared to Lee et al. [48] for photons between 1.9–11.9 GeV, for MSPs within  $10^\circ$  of the Galactic Center and  $|b| > 2^\circ$ . The figure on the right compares to Calore et al. [51] for photons 2–5 GeV from MSPs in the region  $|l| < 10^\circ$  and  $0.5^\circ < |b| < 10^\circ$ . The vertical blue dashed line shows the approximate flux of a source for which  $\sim 1$  photon would have been detected according to Chang et al. [110].

to around the level at which  $\sim 1$  photon would have been detected, where sources would not be distinguishable from diffuse emission [110]. In Fig. 2.20 the MSP flux distribution is generally similar to or below the sub-threshold point source distribution of the more recent work of Calore et al. [51].

As can be seen in Fig. 2.3, the posterior distribution for parameter  $\log_{10}(d_{\text{parallax}}/\text{kpc})$  is limited by the upper limit on the prior. In the limit as  $d_{\text{parallax}} \rightarrow \infty$ , the probability, given by Eq. 2.47, of an MSP having a parallax measurement converges to  $C_{\text{parallax}}$  for all distances  $d$ ; therefore, we imposed the prior upper limit at  $d_{\text{parallax}} = 100$  kpc to avoid this degeneracy. This upper limit will affect the  $\log_{10}(d_{\text{parallax}}/\text{kpc})$  and  $C_{\text{parallax}}$  68% confidence intervals shown in Table 2.3 and Fig. 2.3, however, they are not correlated with any other model parameters.

In order to check the sensitivity of the results to the particular GCE spectrum used we also fitted Model A1 to alternative spectra. Using Bartels et al. [7], we found a small number of changes where the median of a parameter was outside the 68% ranges shown in Table 2.3. The ratio of nuclear bulge and boxy bulge MSPs is lower with  $\log_{10}(N_{\text{nb}}/N_{\text{bb}}) = -0.90^{+0.06}_{-0.06}$ . The relationship between  $\Gamma$  and  $\dot{E}$  is steeper with  $a_{\Gamma} = 0.52^{+0.07}_{-0.08}$ . The fifth DTD bin is constrained strongly to be near 0 with DTD  $p(11.1 - 13.8 \text{ Gyr}) = 0.01^{+0.03}_{-0.01}$ . These changes are likely to be caused by the steep slope on both the low and high energy ends of the boxy bulge spectrum. For the spectrum in Calore et al. [111], we had no nuclear bulge, but otherwise the results were consistent with those using the spectra of Macias et al. [8].

It can be seen in Fig. 2.12 that we find a probability of 0.16 that no bulge MSPs would have been resolved at present, with a median of 2 resolved. With a doubling and quadrupling of detection sensitivity, respectively, a median of 10 and 35 would be expected to be resolved. The detection probability model in Eq. 2.34 models the flux threshold for detection of an MSP as a log-normal distribution. For the median, it uses the Fermi-LAT point source sensitivity at the Galactic coordinates of an MSP  $\mu_{\text{th}}(l, b)$  and multiplies it by a factor of  $10^{K_{\text{th}}}$ , where we find for Model A1,  $K_{\text{th}} = 0.45^{+0.09}_{-0.08}$ . The standard deviation of  $\log_{10}(F_{\text{th}})$  around the median  $F_{\text{th}}$  of  $10^{K_{\text{th}}}\mu_{\text{th}}(l, b)$  is  $\sigma_{\text{th}}$ , which is  $0.28^{+0.05}_{-0.04}$  for Model A1. As we have used MSPs which are not only detected as a point source, but have confirmed gamma-ray pulsations, it is not a surprise that the MSP detection threshold is higher than the point source detection threshold. However, for most of the MSPs, gamma-ray pulsations have been discovered by searching for gamma-ray pulsations from a known radio MSP; therefore, the radio detection sensitivity, which is not explicitly accounted for, may be a significant contributor to the fitted values of  $K_{\text{th}}$  and  $\sigma_{\text{th}}$ . A radio pulse is smeared by dispersion associated with the integrated density of free electrons and by scattering caused by small scale variation in that density, and this may be particularly severe in the region of the Galactic Center [109, 112–114]. As this would lead to

lower sensitivity to radio pulsations, our model may in fact be overestimating the probability of resolving inner Galaxy MSPs. A number of young pulsars have been discovered in the Galactic Center region [115, 116], but the smearing of an MSP pulse will be larger relative to its period. However, the discovery of a magnetar with a period of 3.76 s near Sagittarius A\*, the black hole at the center of the Milky Way, has suggested that the scattering of radio pulses from that region may be small enough that MSPs close to Sagittarius A\* could be detected at radio frequencies  $\gtrsim 10$  GHz [117, 118].

In Table 2.8 we provide a list of resolved MSPs with a significant probability (greater than 5%) of being bulge MSPs according to Model A1. These probabilities are worked out in Eq. B.3 by evaluating the contribution of the bulge to the corresponding MSP's likelihood. The probability of resolved MSP  $i$  being a bulge MSP can be thought of as a draw from a single trial binomial distribution with probability  $p(\text{Bulge MSP}_i)$  and therefore also having an expectation value of  $p(\text{Bulge MSP}_i)$ . Using the linearity of expectation, this implies that the expected number of resolved bulge MSPs in the current data is  $\sum_i p(\text{Bulge MSP}_i) = 1.1$  where the sum is over all resolved MSPs including those listed in Table 2.8. This is at the lower end of, but consistent with, the range seen in Fig. 2.12 produced based on the fitted model parameters.

Note that there is also some systematic uncertainty in the distance to the MSPs which is hard to quantify. For example, in Table 4 of ref. [119] they have a distance of 3.4 kpc for J1747–4036 while we have a distance of  $7.3^{+0.7}_{-0.7}$  kpc. Also in Table 2.2 of ref. [120] they have a distance of 3.1 kpc for J1855–1436 while we have a distance of  $5.3^{+0.6}_{-0.6}$  kpc. These difference may be due to a change in model of the Galactic free electron density. According to the ATNF online database<sup>4</sup> J1747–4036 and J1855–1436 have distances of 7.15 kpc and 5.13 kpc respectively which are compatible with our values. Also, in Table 2 of ref. [121] they have a dispersion measure distance of 1.8 kpc for J1811–2405 while we have a parallax derived distance of  $5^{+11}_{-2}$  kpc.

In Fig. 2.21 the locations of resolved MSPs are shown along with a simulated distribution of disk and bulge MSPs. The elongated nature of the bulge geometry does not play a big role in the probability of having a resolved bulge MSP. This can be seen by changing the boxy bulge geometry in Eq. 2.2 to a spherically symmetric geometry with  $\rho_{\text{boxy bulge}} \propto r^{-2.4}$  up to  $r = 3.1$  kpc and  $\rho_{\text{boxy bulge}} = 0$  for larger radii [15]. We then find that  $2 \pm 2$  MSPs are expected to be resolved for Model A1. Also, for this spherically symmetric bulge case, the MSPs in Table 2.8 have probabilities of 0.5, 0.2, and  $4 \times 10^{-4}$  respectively. So only PSR J1855-1436, with its high  $l = 20.4^\circ$ , is significantly affected by the bulge geometry.

<sup>4</sup><https://www.atnf.csiro.au/research/pulsar/psrcat/>

As shown in Fig. 2.11, the disk, nuclear bulge and boxy bulge have a consistent MSP to stellar mass ratio. This is a good confirmation of our assumption that the population of individual, resolved MSPs belonging mostly to the disk population, on the one hand, and the apparently diffuse  $\gamma$ -ray emission from the GCE, on the other, can be self-consistently explained as arising from MSPs drawn from the same underlying luminosity function given by Eq. 2.24.

Name	Bulge Probability	$l$ (deg)	$b$ (deg)	$d$ (kpc)	$L$ ( $\times 10^{34}$ erg s $^{-1}$ )	$d$ percentile	$L$ percentile
PSR J1747-4036	0.4	-9.8	-6.4	$7.3^{+0.7}_{-0.7}$	$7.7^{+1.8}_{-1.5}$	100	99
PSR J1811-2405	0.5	7.1	-2.5	$5^{+11}_{-2}$	$5^{+47}_{-4}$	96	97
PSR J1855-1436	0.1	20.4	-7.6	$5.3^{+0.6}_{-0.6}$	$1.8^{+0.5}_{-0.4}$	98	91

Table 2.8 Details of MSPs with significant probability (greater than 5%) of coming from bulge population for Model A1. The names are from the Public List of LAT-Detected Gamma-Ray Pulsars.<sup>1</sup> Distances are medians and 68% intervals found by sampling the parallax uncertainty distribution if available or otherwise using the dispersion measure and randomly sampling the parameters of the YMW16 electron density model. Luminosities are medians and 68% intervals found by additionally sampling from the flux uncertainty distribution. Percentiles are the fraction of resolved MSPs with smaller or equal luminosity/distance. This excludes MSPs without available distance estimates.



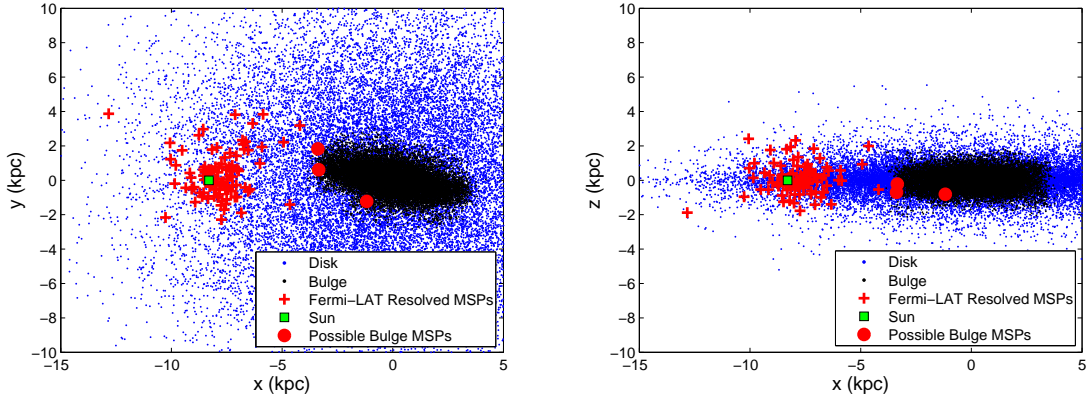


Fig. 2.21 Simulated distribution of disk and bulge MSPs for Model A1 ( $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ ). Also shown are those resolved MSPs with available distance estimates. These figures were made by selecting randomly from the highest likelihood parameter sets in our eight Markov chains for this model. Red circles are the MSPs in Table 2.8.

## 2.5 Conclusion

We compared a wide variety of luminosity function models for the Fermi-LAT gamma-ray MSP data. We found a convincing preference for Model A1 for which  $L = \eta E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$  with a significantly positive  $a_\gamma = 1.2 \pm 0.3$  and  $d_\gamma = 0.5 \pm 0.1$ . Thus we confirm the result obtained by Kalapotharakos et al. [18] that MSP gamma-ray emission is consistent with curvature radiation and inconsistent with synchrotron radiation. By comparing with other models, we showed that the main source of the positive  $a_\gamma$  result was the need to account for a significant logarithmic correlation in the data between the  $E_{\text{cut}}$  and distance in the form of the dispersion measure. We also showed that the main source of the positive  $d_\gamma$  was due to the significant logarithmic correlation in the data between the period and the distance in the form of the dispersion measure.

Additionally, we found that it was warranted to include a relationship between the spectral parameters and  $\dot{E}$ . In particular we found that a linear relationship between the mean of the spectral index  $\mu_\Gamma$  and the  $\log_{10}(\dot{E})$ , as specified in Eq. 2.29, had a significantly positive slope of  $a_\Gamma = 0.4 \pm 0.08$ . We identified the source of this positive slope to be the need to explain the significant correlation between  $\Gamma$  and  $\log_{10}(P)$  seen in the data.

We non-parametrically estimated the delay time distribution of the MSPs but found the current data do not strongly constrain it. Our results obtained using a DTD prescription are also not significantly different from those obtained assuming a uniform age distribution for the MSPs.

Our results demonstrate that the population of MSPs that can explain the gamma-ray signal from the resolved MSPs in the Galactic disk and the unresolved MSPs in the boxy bulge and nuclear bulge can consistently be described as arising from a common evolutionary trajectory for some subset of astrophysical sources common to all these different environments. We do not require that there is anything systematically different about the inner Galaxy MSPs to explain the GCE. We also found that the current data are not accurate enough to be sensitive to the small differences between the bulge and disk MSPs.

We estimated that there are between about 20 and 50 thousand MSPs in the boxy bulge at 68% confidence interval. We identified three candidate resolved MSPs (J1747-4036, J1811-2405, J1855-1436) that have significant probabilities (0.4, 0.5 and 0.1 respectively) of being members of the boxy bulge population. We estimate that this number would increase to 9 and 31 resolved boxy bulge MSPs were the sensitivity to gamma-ray MSPs doubled or quadrupled, respectively.

# Chapter 3

## The Effect of Birth Kicks on the Distribution of Millisecond Pulsars

### 3.1 Introduction

In the “recycling” model of MSP formation a neutron star is spun up to millisecond periods through the transfer of mass from a binary companion. This requires that the binary system survives the kick produced by any asymmetry in the core collapse supernova explosion [56]. However, an alternative to the recycling channel is accretion induced collapse of white dwarfs into neutron stars which may produce more than half of all observed MSPs [54, 122, 123]. In this case the system does not receive a significant natal kick [124, 125]. This would imply that the MSPs have much smaller peculiar velocities in comparison to the recycling model [66, 67, 126].

In Chapter 2, we modelled Fermi-LAT detected MSPs as having a Maxwell distributed peculiar velocity with the scale parameter  $\sigma_v$  found to be  $77 \pm 6 \text{ km s}^{-1}$  where we quote error bars at the 68% confidence interval throughout this chapter. This velocity applies for disk MSPs and we assume it will not be significantly different for bulge MSPs. Although the star formation histories are very different in the bulge and disk, as can be seen from Fig. 2.6 the probability distribution of luminosities in the bulge and disk only differ by a few percent. Also, as can be seen from Fig. 2.11, the bulge and disk have a ratio of number of MSPs formed per solar mass which is within one error bar of each other. Based on this, we assume that the bulge and disk have the same mix of MSP formation channels and thus the same probability distribution of natal kick velocities.

If the source of the GCE is a population of unresolved MSPs, then the spatial distribution may be smoothed to some degree relative to the stellar mass in the bulge. Eckner et al. [127]

used the virial theorem to estimate the “smoothing length” of MSPs as 700 – 900 pc for kicks  $\lesssim 70 \text{ km s}^{-1}$ . However, they assumed a spherically symmetric spatial distribution for the MSPs.

In this chapter we use  $N$ -body simulations to estimate what are the effects of MSP kicks for a boxy bulge distribution. In Section 3.2 we explain our method. Our results are given in Section 3.3, and our conclusions in Section 3.4.

## 3.2 Method

For this work we use the code of Bedorf et al. [128] to run  $N$ -body simulations in order to model the Milky Way.<sup>1</sup> We use parameters corresponding to models MWa, MWb and MWc0.8 as denoted by Fujii et al. [129] as they were the best fitting models to Galaxy observations that Fujii et al. found. Comparing to bulge kinematics, bar length, and pattern speed observations they found  $\chi^2 = 5.3, 8.0, 12.6$  for MWa, MWb, MWc0.8 respectively. For each model we generated a total of 30 million disk, bulge and dark matter halo particles. These initial populations are generated using the methods of Kuijken and Dubinski [130], Widrow and Dubinski [131], and Widrow et al. [132].<sup>2</sup> As in Fujii et al. [129], we use time-steps of  $\sim 0.6 \text{ Myr}$ , an opening angle of 0.4 radians and ran the simulation for 10 Gyr. However, we use a softening length of 30 pc. Also, our dark-matter halo particles have a mass 8 times larger than the disk and bulge particles. Taking into account the masses of the various components, this implies that, out of the 30 million particles, of order 10 million represent stellar mass and the remainder represent dark matter.

In order to model the density of MSPs, we additionally include massless (so they do not affect the simulation) disk and bulge particles that are given a normally distributed perturbation to each component of their velocity vector with mean zero and standard deviation  $\sigma_k$ . The kick velocity magnitude is therefore Maxwell distributed. The probability density function of a Maxwell distribution can be written as:

$$p(x) = \sqrt{\frac{2}{\pi}} \frac{x^2 \exp(-x^2/2\sigma^2)}{\sigma^3} \quad (3.1)$$

where  $x$  is the magnitude of a three dimensional vector with components sampled from the normal distribution  $\mathcal{N}(0, \sigma^2)$ . For each model we try a case where the kicks occurred at the beginning of the  $N$ -body simulations and a case where the kicks occur randomly with a uniform rate over the course of the 10 Gyr.

<sup>1</sup>Available at: <https://github.com/treecode/Bonsai>

<sup>2</sup>We used the implementation at: <https://github.com/treecode/Galactics.parallel>

The first step is to estimate the kick velocity scale required to produce a peculiar velocity distribution consistent with Chapter 2 where for the best model  $\sigma_v = 77 \pm 6 \text{ km s}^{-1}$ . We do this by running each model with 41 populations of  $10^5$  kicked particles with  $\sigma_k$  between 70 and  $110 \text{ km s}^{-1}$ . We separate the velocity of each particle into two components:

$$\mathbf{v} = \mathbf{v}_c + \mathbf{v}_p \quad (3.2)$$

where  $\mathbf{v}_c$  is the velocity of a particle on a circular orbit around the center of the galaxy and  $\mathbf{v}_p$  is the peculiar velocity. The magnitude of  $\mathbf{v}_c$  for a particle with coordinates  $x$ ,  $y$ , and  $z$  can be evaluated using the centripetal force:

$$\|\mathbf{v}_c\| = \sqrt{\|\mathbf{a}_c(x, y, 0)\| R} \quad (3.3)$$

where  $\mathbf{a}_c(x, y, z)$  is the acceleration toward the center of the galaxy and  $R^2 = x^2 + y^2$ . For  $R$  outside the bar region, and for small peculiar velocity, we are therefore assuming that particles are rotating with the disk, with  $\mathbf{v}_c$  the rotation velocity of the disk at  $R$  [133].

We use a maximum likelihood estimate of the final  $\sigma_v$  for each initial  $\sigma_k$ . For a set of  $N$  particles with peculiar velocities  $v_1, \dots, v_N$ , the log-likelihood is obtained by assuming velocities have a Maxwell distribution:

$$\log(L) = \frac{N}{2} \log\left(\frac{2}{\pi}\right) - 3N \log(\sigma_v) + \sum_{i=1}^N 2 \log(v_i) - \frac{v_i^2}{2\sigma_v^2} \quad (3.4)$$

and therefore

$$\frac{d \log(L)}{d \sigma_v} = -\frac{3N}{\sigma_v} + \sum_{i=1}^N \frac{v_i^2}{\sigma_v^3}. \quad (3.5)$$

Then solving for  $\sigma_v$  where  $d \log(L)/d \sigma_v = 0$ , we find the maximum likelihood estimate for  $\sigma_v$  is:

$$\hat{\sigma}_v = \sqrt{\frac{\sum_{i=1}^N v_i^2}{3N}} \quad (3.6)$$

This is done for particles where  $4 \text{ kpc} \leq R \leq 12 \text{ kpc}$  and  $|z| \leq 2 \text{ kpc}$ , ensuring we are estimating the peculiar velocity distribution scale parameter for particles in the disk region from which gamma-ray MSPs are most likely to be resolved and where  $\mathbf{v}_c$  approximately represents disk rotation.

Once we have a best fitting  $\sigma_k$  for each model, we rerun each  $N$ -body simulation using  $2 \times 10^6$  kicked particles with that  $\sigma_k$ . We also have  $2 \times 10^6$  massless particles which are not kicked with the same initial positions and velocities. We then use MCMC to fit a

parametric model to both final particle distributions. This model consists of four components: a spherically symmetric bulge, bar, a long bar and a disk. The spherically symmetric bulge component uses the Hernquist model [134]:

$$\rho_{\text{Hernquist}}(r) \propto \frac{1}{(r/a_b)(1+r/a_b)^3} \quad (3.7)$$

where  $r^2 = x^2 + y^2 + z^2$  and  $a_b$  is a free parameter. The initial conditions used here of Fujii et al. [129] include a component distributed according the Hernquist model. The bar model is distributed as:

$$\rho_{\text{bar}}(R_s) \propto K_0(R_s) \times \begin{cases} 1 & R \leq R_{\text{end}} \\ \exp(-(R - R_{\text{end}})^2/h_{\text{end}}^2) & R > R_{\text{end}} \end{cases} \quad (3.8)$$

where  $K_0$  is the modified Bessel function of the second kind and where:

$$R_{\perp}^{C_{\perp}} = \left(\frac{|x|}{x_b}\right)^{C_{\perp}} + \left(\frac{|y|}{y_b}\right)^{C_{\perp}} \quad (3.9)$$

$$R_s^{C_{\parallel}} = R_{\perp}^{C_{\parallel}} + \left(\frac{|z|}{z_b}\right)^{C_{\parallel}} \quad (3.10)$$

where the free parameters are  $C_{\parallel}$ ,  $C_{\perp}$ ,  $x_b$ ,  $y_b$ ,  $z_b$  and  $R_{\text{end}}$  with  $h_{\text{end}}$  fixed at  $\sqrt{\frac{1}{2}}$  kpc. The effective radius is  $R_s$ ; the scale lengths are  $x_b$ ,  $y_b$ , and  $z_b$ ; and  $C_{\perp}$  and  $C_{\parallel}$  are the face-on and edge-on shape parameters. The bar shape is elliptical in the corresponding direction when  $C_{\perp}, C_{\parallel} = 2$ , diamond-shaped when  $C_{\perp}, C_{\parallel} < 2$ , and boxy when  $C_{\perp}, C_{\parallel} > 2$ . The Gaussian function with scale length  $h_{\text{end}}$  in Eq. (3.8) truncates the bar at radius  $R_{\text{end}}$ . The modified Bessel function was also used in Cao et al. [26] to model the distribution of red clump giants, but with no cutoff and with  $C_{\parallel} = 4$  and  $C_{\perp} = 2$ . For the long bar we use [29]:

$$\rho_{\text{long bar}}(x, y, z) \propto \exp\left(-\left(\left(\frac{|x|}{x_{\text{lb}}}\right)^{C_{\perp, \text{lb}}} + \left(\frac{|y|}{y_{\text{lb}}}\right)^{C_{\perp, \text{lb}}}\right)^{1/C_{\perp, \text{lb}}}\right) \exp\left(-\frac{|z|}{z_{\text{lb}}}\right) \times \text{Cut}\left(\frac{R - R_{\text{out}}}{\sigma_{\text{out}}}\right) \text{Cut}\left(\frac{R_{\text{in}} - R}{\sigma_{\text{in}}}\right) \quad (3.11)$$

where  $C_{\perp,lb}$ ,  $x_{lb}$ ,  $y_{lb}$ ,  $z_{lb}$ ,  $R_{out}$  and  $R_{in}$  are free parameters,  $\sigma_{out} = \sigma_{in} = \sqrt{\frac{1}{2}}$  kpc and:

$$\text{Cut}(x) = \begin{cases} \exp(-x^2) & x > 0 \\ 1 & x \leq 0 \end{cases} \quad (3.12)$$

Finally, we have a disk with a central hole:

$$\rho_{\text{disk}}(x, y, z) \propto \exp(-R^2/2\sigma_r^2) \exp(-|z|/z_0) H(x, y) \quad (3.13)$$

where  $\sigma_r$  and  $z_0$  are free parameters and for the hole we use the form adopted by Freudenreich [1]:

$$H(x, y) = 1 - \exp\left(-(R_H/O_R)^{O_N}\right) \quad (3.14)$$

with:

$$R_H^2 = (x)^2 + (\epsilon y)^2 \quad (3.15)$$

where  $\epsilon$ ,  $O_R$  and  $O_N$  are also free parameters.

The total number of particles in our simulations are fixed. So we do not have to include the number of particles as part of our likelihood. Therefore the probability of having an  $N$ -body particle at position  $x, y, z$  will be proportional to the density of our model ( $\rho$ ) at  $x, y, z$ . We have for each component of the model a parameter giving the probability a particle is from that component. We treat the probability of an  $N$ -body particle being from a component of the density distribution as parameters. These parameters,  $P(\text{Disk})$ ,  $P(\text{Bar})$ ,  $P(\text{Hernquist})$  and  $P(\text{Long Bar})$ , have a Dirichlet prior [108]. This prior constrains

$$P(\text{Disk}) + P(\text{Bar}) + P(\text{Hernquist}) + P(\text{Long Bar}) = 1$$

and is uniformly distributed over any values of these parameter satisfying that condition. The likelihood is then:

$$\log(L) = \sum_i^N \log(\rho(x_i, y_i, z_i)) \quad (3.16)$$

where  $x_i$ ,  $y_i$  and  $z_i$  are the coordinates of a particle,  $N$  is the number of particles, and  $\rho$  is the density of the model:

$$\begin{aligned} \rho(x, y, z) = & P(\text{Disk})\rho_{\text{disk}}(x, y, z) + P(\text{Bar})\rho_{\text{bar}}(x, y, z) \\ & + P(\text{Hernquist})\rho_{\text{Hernquist}}(x, y, z) + P(\text{Long Bar})\rho_{\text{long bar}}(x, y, z) \end{aligned} \quad (3.17)$$

All scale parameters are given a prior so they are uniform in  $\log(\theta)$  where  $\theta \in \{a_b, x_b, y_b, z_b, x_{lb}, y_{lb}, z_{lb}, \sigma_r, z_0, O_R\}$  and this implies  $p(\theta) \propto 1/\theta$ . In calculating the likelihood, we do not include particles for which  $R > 12$  kpc or  $|z| > 3$ . It can be seen in Fujii et al. [129] that the scale height of the disk may start to decline between  $10 \lesssim R \lesssim 15$  kpc. We also don't want the fit to be affected by particles that may have been kicked well out of the galaxy. The likelihood ( $L$ ) is insensitive to being multiplied by a constant but that constant has to be the same for all parameters of our combined model. To accommodate this we normalize each density component such that

$$\int_{R \leq 12 \text{ kpc}, |z| \leq 3 \text{ kpc}} \rho_i(x, y, z) dx dy dz = 1$$

where  $i \in \{\text{disk, bar, Hernquist, long bar}\}$ . This integral is estimated with importance sampling. We use a set of random numbers which are transformed into the points at which we evaluate the density models in order to estimate the normalization constant. In order to stabilize the estimation of the likelihood function these numbers are always the same every time we perform the importance sampling within a particular chain.

After running the  $N$ -body simulations, we shift the coordinates of the particles so that the center of mass is at the origin, then rotate so the bar is along the  $x$ -axis. The bar angle is estimated using the method described in Fujii et al. [129]. However, we add four parameters that we expect to be near zero to allow a further shift in the center and clockwise rotation of the model. These are  $\alpha$ ,  $x_{\text{center}}$ ,  $y_{\text{center}}$  and  $z_{\text{center}}$ , with the latter three parameters in parsecs, so:

$$\begin{aligned} x_{\text{data}} &= \cos(\alpha)x + \sin(\alpha)y + x_{\text{center}}/1000 \\ y_{\text{data}} &= -\sin(\alpha)x + \cos(\alpha)y + y_{\text{center}}/1000 \\ z_{\text{data}} &= z + z_{\text{center}}/1000 \end{aligned} \tag{3.18}$$

where  $x_{\text{data}}$ ,  $y_{\text{data}}$  and  $z_{\text{data}}$  are coordinates in the coordinate system of the  $N$ -body simulation. In estimating the peculiar velocity distribution scale parameter,  $\hat{\sigma}_v$ , above, we assumed  $x \approx x_{\text{data}}$ ,  $y \approx y_{\text{data}}$  and  $z \approx z_{\text{data}}$ .

In Chapter 2 we used for MCMC the adaptive Metropolis algorithm of Haario et al. [98], here, however, we found it was necessary to replace this with an alternative algorithm to ensure rapid convergence to the peak likelihood region of the parameter space. The MCMC algorithm used in this chapter is similar to that of Foreman-Mackay et al. [135] with a mixture of the Differential Evolution [136] and snooker updates [137]. Instead of performing a single random walk through the parameter space where proposed moves are accepted with the



probability  $r$  of Eq. 2.35, we use an ensemble of  $K$  “walkers” where a proposed update for a walker  $j$  depends on the distribution of the other walkers.

A single step of the stretch move update suggested in Foreman-Mackay et al. [135] involves updating the  $K$  walkers sequentially. Let  $\mathbf{x}_j$  be the state of walker  $j$ , an update for walker  $\mathbf{x}_j$  is performed as follows:

1. Draw  $k$  from  $1, 2, \dots, K$  where  $k \neq j$
2. Draw  $z$  from the probability density function with parameter  $a$  (Foreman-Mackay et al. [135] suggest  $a = 2$ ):

$$g(z) \propto \begin{cases} \frac{1}{\sqrt{z}} & z \in [\frac{1}{a}, a] \\ 0 & \text{otherwise} \end{cases} \quad (3.19)$$

3. Calculate proposal  $\mathbf{y} = \mathbf{x}_k + z(\mathbf{x}_j - \mathbf{x}_k)$
4. Calculate acceptance probability  $r$ :

$$r = z^{d-1} \frac{p(\mathbf{y})}{p(\mathbf{x}_j)} \quad (3.20)$$

where  $d$  is the number of dimensions

5. Set  $\mathbf{x}_j = \mathbf{y}$  with probability  $\min(1, r)$

We found better results using a mixture of two alternative updates: 80% the Differential Evolution update of ter Braak [136] and 20% the snooker update of ter Braak and Vrugt [137].

To update walker  $\mathbf{x}_j$  using the Differential Evolution update:

1. Draw  $k$  and  $l$  from  $1, 2, \dots, K$  where  $k \neq j$ ,  $l \neq j$  and  $l \neq k$
2. Propose  $\mathbf{y} = \mathbf{x}_j + \gamma(\mathbf{x}_k - \mathbf{x}_l) + \mathbf{e}$  where  $\gamma$  is a parameter and where  $\mathbf{e}$  is drawn from a small  $d$  dimensional symmetric probability distribution
3. Calculate acceptance probability  $r$ :

$$r = \frac{p(\mathbf{y})}{p(\mathbf{x}_j)} \quad (3.21)$$

4. Set  $\mathbf{x}_j = \mathbf{y}$  with probability  $\min(1, r)$

We drew  $\mathbf{e}$  from a  $d$  dimensional Gaussian with standard deviation  $10^{-5}$  in each dimension. In case the likelihood distribution had multiple modes, we used  $\gamma = 1$  with probability 0.1 as suggested by ter Braak [136], otherwise we used the default value of  $\gamma = 2.38/\sqrt{2d}$ .

Using the snooker update, we update  $\mathbf{x}_j$  as follows:

1. Draw  $k, l$  and  $m$  from  $1, 2, \dots, K$  with no index repeated or equal to  $j$
2. Calculate the orthogonal projections of  $\mathbf{x}_l$  and  $\mathbf{x}_m$  onto the line  $\mathbf{x}_j - \mathbf{x}_k$ ,  $\text{proj}_{\mathbf{x}_j - \mathbf{x}_k}(\mathbf{x}_l)$  and  $\text{proj}_{\mathbf{x}_j - \mathbf{x}_k}(\mathbf{x}_m)$ , where:

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} \quad (3.22)$$

3. Propose  $\mathbf{y} = \mathbf{x}_j + \gamma_s \left( \text{proj}_{\mathbf{x}_j - \mathbf{x}_k}(\mathbf{x}_l) - \text{proj}_{\mathbf{x}_j - \mathbf{x}_k}(\mathbf{x}_m) \right)$  where  $\gamma_s$  is a parameter
4. Calculate acceptance probability  $r$ :

$$r = \frac{p(\mathbf{y}) |\mathbf{y} - \mathbf{x}_k|^{d-1}}{p(\mathbf{x}_j) |\mathbf{x}_j - \mathbf{x}_k|^{d-1}} \quad (3.23)$$

5. Set  $\mathbf{x}_j = \mathbf{y}$  with probability  $\min(1, r)$

We use  $\gamma_s = 2.38/\sqrt{2}$  as suggested by ter Braak and Vrugt [137].

We also used a simple annealing method in which we divide the log-likelihood by a temperature  $T$  which is gradually reduced to 1. The posterior probability density for a parameter set  $\theta$  at MCMC iteration  $t$  is:

$$p(\theta | N\text{-body data}) \propto p(\theta) L^{1/T_t} \quad (3.24)$$

where  $p(\theta)$  is the prior,  $L$  is the likelihood, and  $T_t$  is the temperature. We used a linearly decreasing  $\log(T)$  from  $\log(1000)$  to  $\log(1)$  during the first half of each Markov chain, which we discard. This method allows the algorithm to explore a broad region in the parameter space, while slowly converging to the desired posterior distribution where  $T = 1$ . This appeared to help the Markov chains avoid getting stuck in local likelihood maxima.

### 3.3 Results

In Table 3.1 we present the kick velocities  $\sigma_k$  that produce peculiar velocity distributions close to  $\sigma_v = 77 \pm 6 \text{ km s}^{-1}$  as estimated in Chapter 2. We display the rotation curves at  $t = 10 \text{ Gyr}$  for the three models in Fig. 3.1. The central values in Table 3.1 were used to

run  $N$ -body simulations with a larger number of particles to which we fitted a parametric model. The fitted parameters are shown in Tables 3.2, 3.3 and 3.4. There are three potential sources of uncertainty in the model parameters: the posterior density function, variation in the likelihood between chains as a result of the importance sampling method used to estimate the normalization constant for each model component, and the possibility that Markov chains may get stuck in different local likelihood maxima. We found that for most parameters the posterior distributions overlapped significantly or, in many cases, were indistinguishable. For a few parameters we had outlier chains; these can be seen in the tables as parameters with large, highly asymmetric uncertainties. For the kicked distributions, the disk model hole parameters were very uncertain with different Markov chains settling on a wide range of different values. As this hole tended to be smaller and/or less sharp, we simply removed it by setting  $H(x,y) = 1$ . We did find, including prior to removing the  $\rho_{\text{disk}}$  hole, the long bar model often acts like a second disk component when fitting to the kicked particles, this disk-like long bar was still allowed a hole through the parameter  $R_{\text{in}}$ .

We ran the  $N$ -body simulations for the two kick rate scenarios separately, but we combine the two sets of Markov chains for the “No Kick” columns in the parameter tables. Typically, the posterior distributions were very similar for Markov chains generated using these two sets of data, slightly expanding our 68% intervals. However, we found a very significant difference was the location of the center. For this reason, we only report the change in parameters  $\alpha$ ,  $x_{\text{center}}$ ,  $y_{\text{center}}$  and  $z_{\text{center}}$  for the fits to the kicked distributions by subtracting off the median of the corresponding Markov chains. For  $\alpha$  the fitted 68% intervals for no kick were  $-2^\circ$  to  $-1^\circ$  for all models,  $|x_{\text{center}}|$  and  $|y_{\text{center}}|$  were always  $\lesssim 40$  pc, and  $|z_{\text{center}}|$  was  $< 2$  pc.

We show the density maps for MWa, MWb and MWc0.8 in Figs. 3.2, 3.3 and 3.4 respectively. These figures include both the  $N$ -body particle data as well as the fitted models. These density maps were produced by binning particles (either the  $N$ -body particles or particles drawn from the fitted model) within 0.25 kpc of the  $x$ - $y$ ,  $x$ - $z$  and  $y$ - $z$  planes. The density maps for the fitted models were generated by taking the average in each bin for simulated data generated using 500 random parameter sets from our Markov chains in each case. In Fig. 3.5 we show the  $N$ -body simulation data and simulated particle distributions along the  $x$ ,  $y$  and  $z$  axes for the MWa case with the kicks occurring at the beginning. The corresponding uniform kick rate case is displayed in Fig. 3.6. The MWb and MWc0.8 cases are shown in Appendix E as Figs. E.1, E.2, E.3 and E.4. In these figures, we bin all particles within 0.25 kpc in the two perpendicular axes. For MWa, we display in Fig. 3.7 the integrated flux along lines of sight in the central  $50^\circ \times 50^\circ$  of the galaxy. The Sun is placed at a distance of 7.9 kpc, at an angle relative to the bar of  $20^\circ$  and at a height of 15 pc [10]. These figures

	Kick At Beginning	Uniform Kick Rate
MWa	$93 \pm 10$	$85 \pm 9$
MWb	$97 \pm 10$	$84 \pm 10$
MWc0.8	$95 \pm 10$	$84 \pm 9$

Table 3.1 Kick velocity Maxwell distribution (Eq. (3.1)) parameters ( $\sigma_k$  in km/s) that produce a peculiar velocity distribution where  $\sigma_v = 77 \pm 6$  km/s.

were generated by binning particles in Galactic latitude and longitude with weights of  $1/d^2$  where  $d$  is the distance of a particle from the Sun. We excluded particles within 1 kpc of the Sun to reduce noise. The corresponding figures for MWb and MWc0.8 are shown in Appendix E as Figs. E.5 and E.6.

In Table 3.5 we show the change in  $-2\log(L)$  when replacing  $\rho_{\text{bar}}(R_s)$  with a range of different models from Freudenreich [1] and Cao et al. [26]. As there is variation in the likelihood between chains, we also show the standard deviation in  $-2\log(L)$ . Our choice of  $\rho_{\text{bar}}(R_s) \propto K_0(R_s)$  is clearly preferred over the others. The worst form, where  $\rho_{\text{bar}}(R_s) \propto \exp(R_s^{-n})$ , was entirely removed with  $P(\text{Bar}) = 0$  and the long bar component taking over the fit in the central region.

In order to estimate the kick effects on the Galactic bulge we used a linear fit to our simulation results of the form

$$\theta_{\text{kicked},i} = \alpha_i \theta_i + \beta_i \quad (3.25)$$

where  $\theta_{\text{kicked},i}$  are the bulge parameters,  $x_b$ ,  $y_b$ ,  $z_b$ ,  $C_\perp$ , and  $C_\parallel$  for the kicked distribution and  $\theta_i$  are the corresponding parameters in the non-kicked case. The  $\alpha_i$  and  $\beta_i$  were found by performing a least squared fit for the values given in Tables 3.2, 3.3, and 3.4. The results are shown in Table 3.6 and Fig. 3.8. A prediction for the Milky Way bulge parameters found in ref. [26] are shown in Table 3.7. The predicted line of sight contours for the kicked and unkicked Milky Way bulge are shown in Fig. 3.9.

### 3.4 Discussion and Conclusions

Our goal was to investigate the effect of neutron star birth kicks on the distribution of MSPs in the Galactic Center. We began by running  $N$ -body simulations with small populations of particles kicked with a range of scales in order to estimate the required Maxwellian kick to produce a peculiar velocity distribution similar to that of resolved gamma-ray MSPs. We

Parameter	No Kick	Kick At Beginning	Uniform Kick Rate
$P(\text{Disk})$	$0.499^{+0.006}_{-0.005}$	$0.418^{+0.006}_{-0.006}$	$0.365^{+0.006}_{-0.004}$
$P(\text{Bar})$	$0.354^{+0.004}_{-0.005}$	$0.3220^{+0.0019}_{-0.0021}$	$0.364^{+0.005}_{-0.007}$
$P(\text{Hernquist})$	$0.0145^{+0.0013}_{-0.0018}$	$0.059^{+0.003}_{-0.002}$	$0.0647^{+0.0022}_{-0.0024}$
$P(\text{Long Bar})$	$0.1323^{+0.0025}_{-0.0025}$	$0.201^{+0.007}_{-0.007}$	$0.206^{+0.007}_{-0.010}$
$\sigma_r$ (kpc)	$4.92^{+0.04}_{-0.03}$	$5.66^{+0.05}_{-0.04}$	$6.07^{+0.05}_{-0.08}$
$z_0$ (kpc)	$0.2238^{+0.0008}_{-0.0027}$	$1.203^{+0.023}_{-0.013}$	$1.222^{+0.018}_{-0.004}$
$O_R$ (kpc)	$2.82^{+0.08}_{-0.11}$	—	—
$O_N$	$4.3^{+0.3}_{-0.2}$	—	—
$\varepsilon$	$0.780^{+0.020}_{-0.013}$	—	—
$a_b$ (kpc)	$0.205^{+0.009}_{-0.010}$	$0.509^{+0.019}_{-0.010}$	$0.555^{+0.022}_{-0.021}$
$C_\perp$	$1.84^{+0.02}_{-0.05}$	$1.831^{+0.013}_{-0.015}$	$1.821^{+0.017}_{-0.015}$
$C_\parallel$	$3.08^{+0.11}_{-0.11}$	$2.49^{+0.03}_{-0.03}$	$2.542^{+0.019}_{-0.019}$
$x_b$ (kpc)	$0.557^{+0.008}_{-0.008}$	$0.685^{+0.005}_{-0.005}$	$0.673^{+0.007}_{-0.011}$
$y_b$ (kpc)	$0.367^{+0.006}_{-0.003}$	$0.455^{+0.003}_{-0.003}$	$0.445^{+0.006}_{-0.011}$
$z_b$ (kpc)	$0.2557^{+0.0014}_{-0.0010}$	$0.3328^{+0.0014}_{-0.0015}$	$0.3155^{+0.0026}_{-0.0029}$
$R_{\text{end}}$ (kpc)	$2.00^{+0.05}_{-0.06}$	$4.6^{+0.4}_{-0.3}$	$4.9^{+0.6}_{-0.6}$
$\Delta\alpha$ (deg)	—	$-0.34^{+0.15}_{-0.14}$	$-0.07^{+0.13}_{-0.13}$
$\Delta x_{\text{center}}$ (pc)	—	$0.8^{+1.0}_{-1.1}$	$-0.7^{+1.0}_{-1.0}$
$\Delta y_{\text{center}}$ (pc)	—	$-1.8^{+0.9}_{-0.9}$	$-0.3^{+0.7}_{-0.7}$
$\Delta z_{\text{center}}$ (pc)	—	$-0.9^{+0.6}_{-0.5}$	$2.7^{+0.4}_{-0.5}$
$x_{\text{lb}}$ (kpc)	$5^{+5}_{-1}$	$2.78^{+0.03}_{-0.03}$	$2.46^{+0.04}_{-0.03}$
$y_{\text{lb}}$ (kpc)	$1.43^{+0.16}_{-0.07}$	$2.78^{+0.04}_{-0.04}$	$2.49^{+0.05}_{-0.04}$
$z_{\text{lb}}$ (kpc)	$0.315^{+0.029}_{-0.010}$	$0.541^{+0.006}_{-0.011}$	$0.495^{+0.010}_{-0.014}$
$C_{\perp,\text{lb}}$	$0.88^{+0.04}_{-0.11}$	$1.764^{+0.026}_{-0.025}$	$1.983^{+0.028}_{-0.026}$
$R_{\text{out}}$ (kpc)	$2.86^{+0.04}_{-0.05}$	$7.46^{+0.06}_{-0.04}$	$8.37^{+0.06}_{-0.06}$
$R_{\text{in}}$ (kpc)	$1.676^{+0.018}_{-0.016}$	$2.09^{+0.03}_{-0.04}$	$2.03^{+0.07}_{-0.12}$

Table 3.2 Fitted parameters for model MWa. We have used the median of the MCMC chains for the central value and also included 68% confidence intervals.

Parameter	No Kick	Kick At Beginning	Uniform Kick Rate
$P(\text{Disk})$	$0.506^{+0.010}_{-0.005}$	$0.6144^{+0.0024}_{-0.0026}$	$0.36^{+0.03}_{-0.04}$
$P(\text{Bar})$	$0.322^{+0.005}_{-0.003}$	$0.244^{+0.007}_{-0.003}$	$0.317^{+0.006}_{-0.003}$
$P(\text{Hernquist})$	$0.0140^{+0.0012}_{-0.0009}$	$0.0470^{+0.0022}_{-0.0021}$	$0.0819^{+0.0017}_{-0.0016}$
$P(\text{Long Bar})$	$0.154^{+0.012}_{-0.010}$	$0.094^{+0.004}_{-0.006}$	$0.24^{+0.04}_{-0.04}$
$\sigma_r$ (kpc)	$5.64^{+0.06}_{-0.07}$	$5.36^{+0.03}_{-0.03}$	$6.9^{+0.7}_{-0.3}$
$z_0$ (kpc)	$0.2249^{+0.0003}_{-0.0003}$	$1.116^{+0.004}_{-0.014}$	$1.30^{+0.05}_{-0.06}$
$O_R$ (kpc)	$3.19^{+0.05}_{-0.09}$	—	—
$O_N$	$4.25^{+0.09}_{-0.17}$	—	—
$\epsilon$	$0.762^{+0.019}_{-0.019}$	—	—
$a_b$ (kpc)	$0.202^{+0.004}_{-0.005}$	$0.585^{+0.023}_{-0.025}$	$0.614^{+0.014}_{-0.011}$
$C_{\perp}$	$2.04^{+0.08}_{-0.10}$	$2.02^{+0.03}_{-0.03}$	$1.873^{+0.016}_{-0.017}$
$C_{\parallel}$	$3.85^{+0.14}_{-0.08}$	$3.06^{+0.05}_{-0.05}$	$2.90^{+0.04}_{-0.04}$
$x_b$ (kpc)	$0.548^{+0.029}_{-0.009}$	$0.587^{+0.026}_{-0.009}$	$0.710^{+0.010}_{-0.006}$
$y_b$ (kpc)	$0.339^{+0.006}_{-0.005}$	$0.406^{+0.013}_{-0.005}$	$0.411^{+0.010}_{-0.005}$
$z_b$ (kpc)	$0.243^{+0.004}_{-0.002}$	$0.306^{+0.007}_{-0.003}$	$0.303^{+0.004}_{-0.002}$
$R_{\text{end}}$ (kpc)	$2.53^{+0.10}_{-0.15}$	$3.37^{+0.10}_{-0.08}$	$5.2^{+0.6}_{-0.4}$
$\Delta\alpha$ (deg)	—	$-0.04^{+0.13}_{-0.13}$	$0.30^{+0.11}_{-0.12}$
$\Delta x_{\text{center}}$ (pc)	—	$4.4^{+1.2}_{-1.2}$	$2.5^{+1.2}_{-1.3}$
$\Delta y_{\text{center}}$ (pc)	—	$1.8^{+0.9}_{-0.8}$	$-1.2^{+0.9}_{-0.8}$
$\Delta z_{\text{center}}$ (pc)	—	$-0.6^{+0.7}_{-0.8}$	$1.4^{+0.5}_{-0.5}$
$x_{\text{lb}}$ (kpc)	$2.7^{+0.8}_{-0.3}$	$2.1^{+0.3}_{-0.1}$	$2.78^{+0.04}_{-0.02}$
$y_{\text{lb}}$ (kpc)	$1.14^{+0.17}_{-0.05}$	$0.93^{+0.08}_{-0.03}$	$2.65^{+0.07}_{-0.04}$
$z_{\text{lb}}$ (kpc)	$0.414^{+0.006}_{-0.016}$	$0.656^{+0.011}_{-0.020}$	$0.56^{+0.04}_{-0.03}$
$C_{\perp, \text{lb}}$	$0.97^{+0.03}_{-0.11}$	$1.61^{+0.04}_{-0.04}$	$1.86^{+0.03}_{-0.03}$
$R_{\text{out}}$ (kpc)	$3.66^{+0.22}_{-0.11}$	$4.67^{+0.05}_{-0.05}$	$9.32^{+0.07}_{-0.25}$
$R_{\text{in}}$ (kpc)	$1.69^{+0.04}_{-0.03}$	$1.78^{+0.04}_{-0.03}$	$1.71^{+0.10}_{-0.06}$

Table 3.3 Fitted parameters for model MWb.

Parameter	No Kick	Kick At Beginning	Uniform Kick Rate
$P(\text{Disk})$	$0.555^{+0.003}_{-0.009}$	$0.467^{+0.012}_{-0.010}$	$0.357^{+0.013}_{-0.003}$
$P(\text{Bar})$	$0.318^{+0.004}_{-0.005}$	$0.262^{+0.003}_{-0.005}$	$0.305^{+0.003}_{-0.002}$
$P(\text{Hernquist})$	$0.0099^{+0.0006}_{-0.0005}$	$0.0603^{+0.0021}_{-0.0025}$	$0.0619^{+0.0019}_{-0.0020}$
$P(\text{Long Bar})$	$0.116^{+0.014}_{-0.004}$	$0.210^{+0.014}_{-0.010}$	$0.277^{+0.005}_{-0.017}$
$\sigma_r$ (kpc)	$5.49^{+0.03}_{-0.04}$	$6.11^{+0.07}_{-0.09}$	$6.99^{+0.06}_{-0.19}$
$z_0$ (kpc)	$0.2246^{+0.0007}_{-0.0022}$	$1.266^{+0.016}_{-0.028}$	$1.262^{+0.026}_{-0.022}$
$O_R$ (kpc)	$2.5^{+0.4}_{-0.16}$	—	—
$O_N$	$4.9^{+2.6}_{-1.1}$	—	—
$\varepsilon$	$0.82^{+0.12}_{-0.12}$	—	—
$a_b$ (kpc)	$0.220^{+0.006}_{-0.006}$	$0.72^{+0.01}_{-0.04}$	$0.74^{+0.03}_{-0.03}$
$C_\perp$	$1.88^{+0.03}_{-0.03}$	$1.861^{+0.020}_{-0.020}$	$1.909^{+0.024}_{-0.022}$
$C_\parallel$	$3.37^{+0.05}_{-0.05}$	$2.69^{+0.03}_{-0.04}$	$2.69^{+0.03}_{-0.03}$
$x_b$ (kpc)	$0.551^{+0.012}_{-0.014}$	$0.657^{+0.006}_{-0.007}$	$0.620^{+0.006}_{-0.005}$
$y_b$ (kpc)	$0.342^{+0.007}_{-0.008}$	$0.434^{+0.005}_{-0.007}$	$0.378^{+0.005}_{-0.003}$
$z_b$ (kpc)	$0.2395^{+0.0015}_{-0.0018}$	$0.3217^{+0.0023}_{-0.0023}$	$0.2845^{+0.0025}_{-0.0018}$
$R_{\text{end}}$ (kpc)	$2.03^{+0.07}_{-0.10}$	$4.81^{+0.29}_{-0.25}$	$5.5^{+2.1}_{-0.9}$
$\Delta\alpha$ (deg)	—	$-0.79^{+0.18}_{-0.18}$	$0.32^{+0.13}_{-0.13}$
$\Delta x_{\text{center}}$ (pc)	—	$1.4^{+1.4}_{-1.3}$	$1.0^{+1.2}_{-1.1}$
$\Delta y_{\text{center}}$ (pc)	—	$-0.2^{+0.9}_{-0.9}$	$0.5^{+0.8}_{-0.8}$
$\Delta z_{\text{center}}$ (pc)	—	$2.4^{+0.7}_{-0.6}$	$-0.5^{+0.5}_{-0.5}$
$x_{\text{lb}}$ (kpc)	$3^{+7}_{-1}$	$2.533^{+0.023}_{-0.022}$	$2.533^{+0.020}_{-0.021}$
$y_{\text{lb}}$ (kpc)	$1.5^{+0.4}_{-0.2}$	$2.584^{+0.027}_{-0.025}$	$2.417^{+0.021}_{-0.022}$
$z_{\text{lb}}$ (kpc)	$0.335^{+0.006}_{-0.005}$	$0.564^{+0.019}_{-0.006}$	$0.530^{+0.010}_{-0.011}$
$C_{\perp, \text{lb}}$	$0.95^{+0.07}_{-0.11}$	$1.964^{+0.029}_{-0.029}$	$1.954^{+0.022}_{-0.022}$
$R_{\text{out}}$ (kpc)	$3.0^{+0.1}_{-0.4}$	$7.88^{+0.07}_{-0.09}$	$9.13^{+0.13}_{-0.08}$
$R_{\text{in}}$ (kpc)	$1.61^{+0.03}_{-0.06}$	$1.91^{+0.05}_{-0.07}$	$1.60^{+0.05}_{-0.03}$

Table 3.4 Fitted parameters for model MWc0.8.

$\rho_{\text{bar}}(R_s)$	Mean $-2(\Delta \log(L))$	Standard Deviation
$K_0(R_s)$	0	739
$\exp(-R_s)$	3581	1606
$\text{sech}^2(R_s)$	12795	1985
$\exp(-0.5R_s^2)$	19369	391
$(1 + R_s^n)^{-1}$	20403	953
$\exp(R_s^{-n})$	57114	3152

Table 3.5 Change in mean  $-2\log(L)$  using different bar models  $\rho_{\text{bar}}(R_s)$ . The mean was taken over all the samples in the MCMC chains.

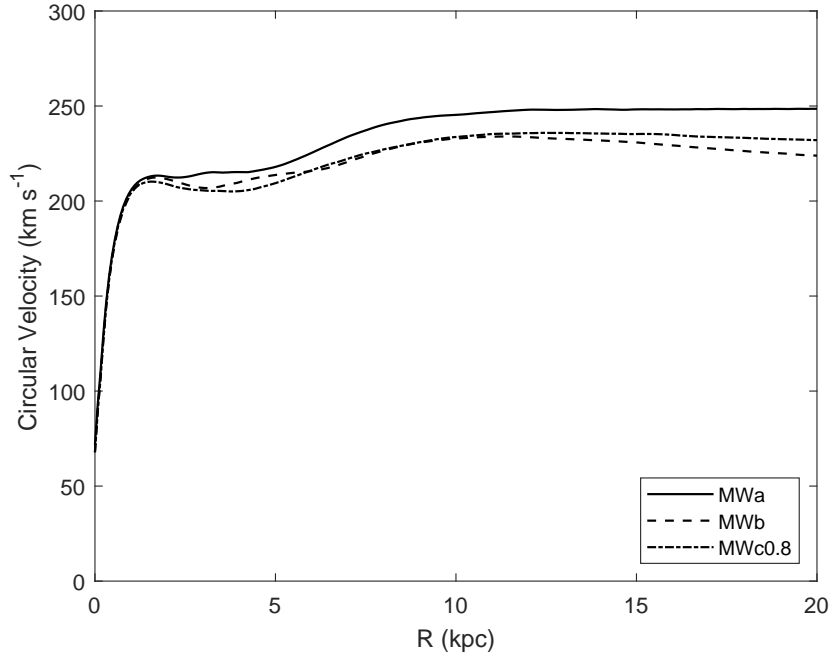


Fig. 3.1 The rotation curves for the three models MWa, MWb and MWc0.8. The local circular velocity of the Sun is  $238 \pm 15 \text{ km s}^{-1}$  [107]. The distance between the Sun and the Galactic Center is approximately 8 kpc.

Parameter	$\alpha_i$	$\beta_i$
$x_b, y_b, z_b$	$1.13 \pm 0.05$	$0.03 \pm 0.02$
$C_{\perp}, C_{\parallel}$	$0.56 \pm 0.02$	$0.81 \pm 0.05$

Table 3.6 Least square fit values with 68% confidence intervals for Eq. 3.25 fitted to the points shown in Figure 3.8.

Parameter	$x_b$ (kpc)	$y_b$ (kpc)	$z_b$ (kpc)	$C_{\perp}$	$C_{\parallel}$
Not kicked	0.67	0.29	0.27	2	4
Kicked	$0.79 \pm 0.02$	$0.36 \pm 0.01$	$0.35 \pm 0.01$	$1.93 \pm 0.02$	$3.05 \pm 0.03$

Table 3.7 Predictions with 68% confidence intervals for the kicked spatial distribution for the Milky Way bulge model found in ref. [26] using Eq. 3.25 and the parameter values given in Table 3.6.



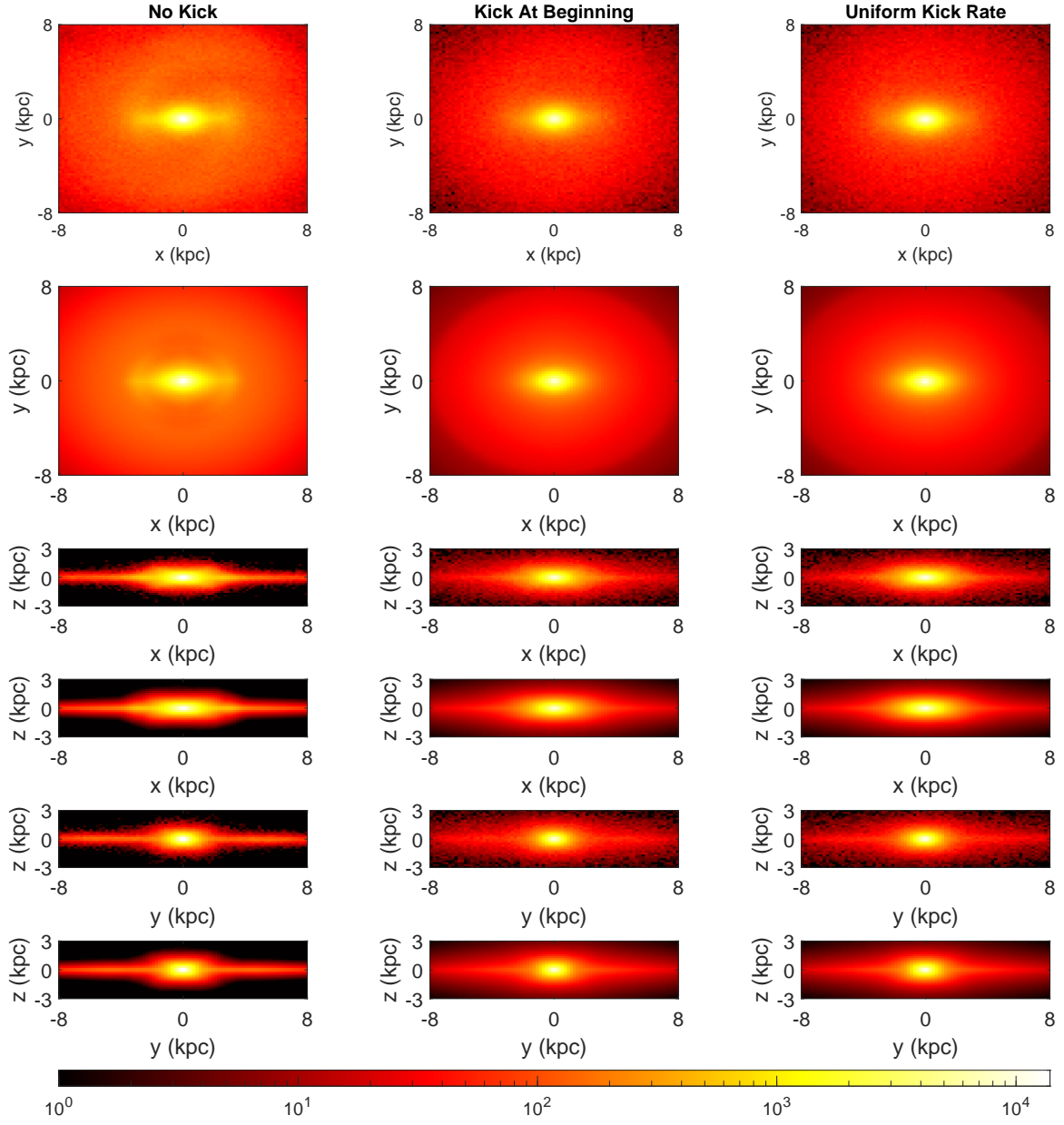


Fig. 3.2 Density map of particles with no kick, a kick at the beginning and a uniform kick rate for the model MWa. Every second row shows the fitted model.

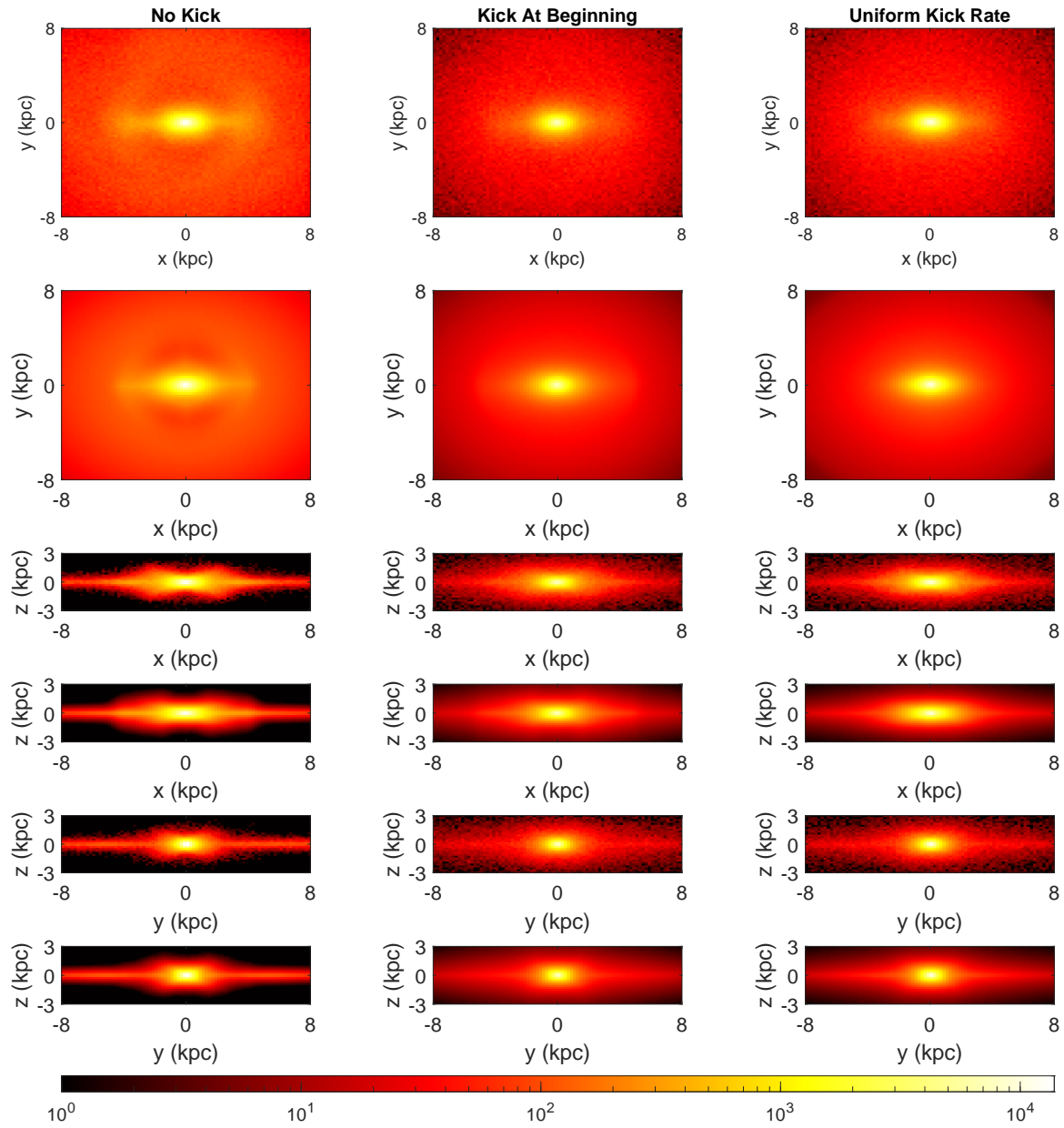


Fig. 3.3 Density map of particles with no kick, a kick at the beginning and a uniform kick rate for the model MWb. Every second row shows the fitted model.

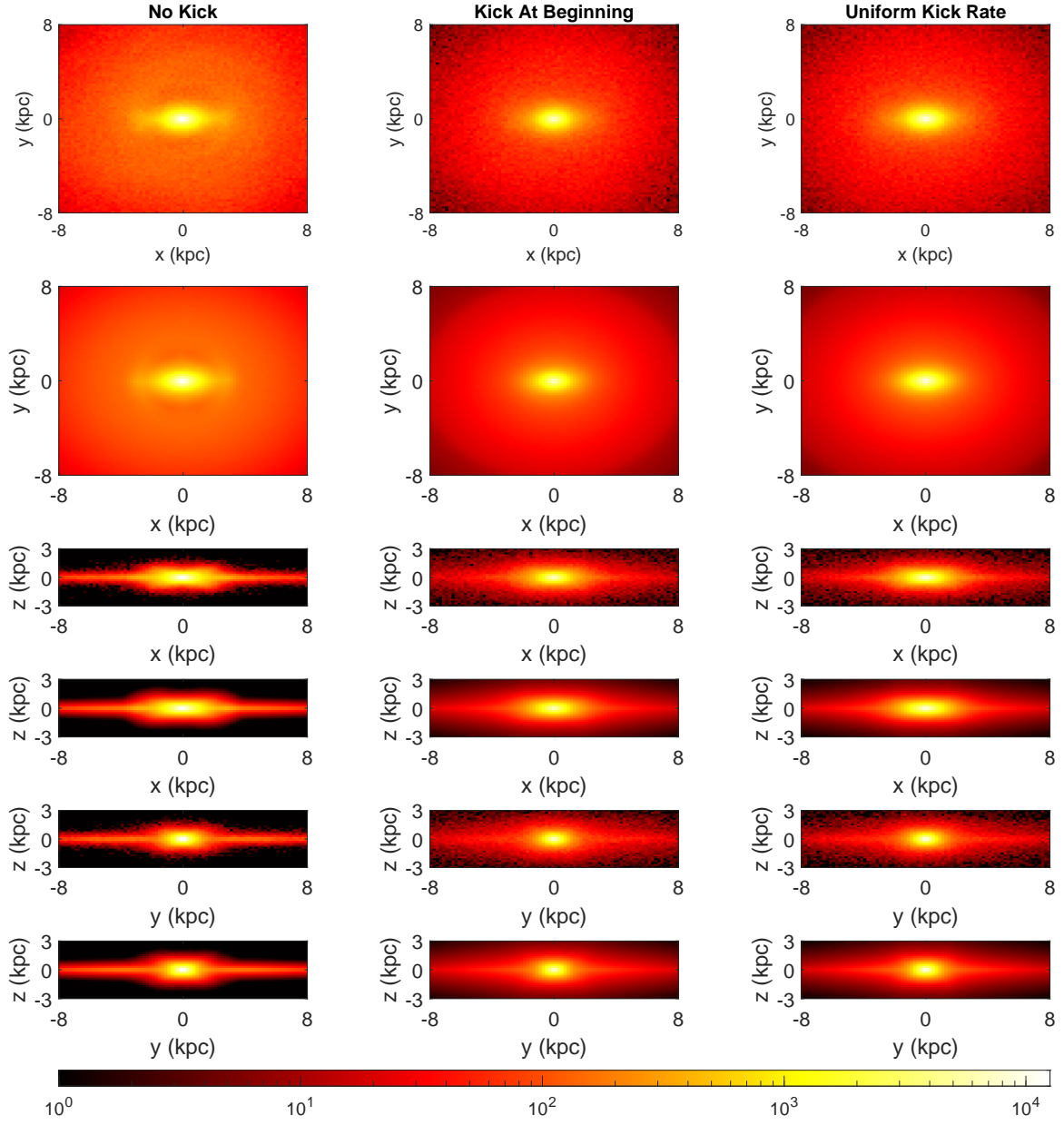


Fig. 3.4 Density map of particles with no kick, a kick at the beginning and a uniform kick rate for the model MWc0.8. Every second row shows the fitted model.

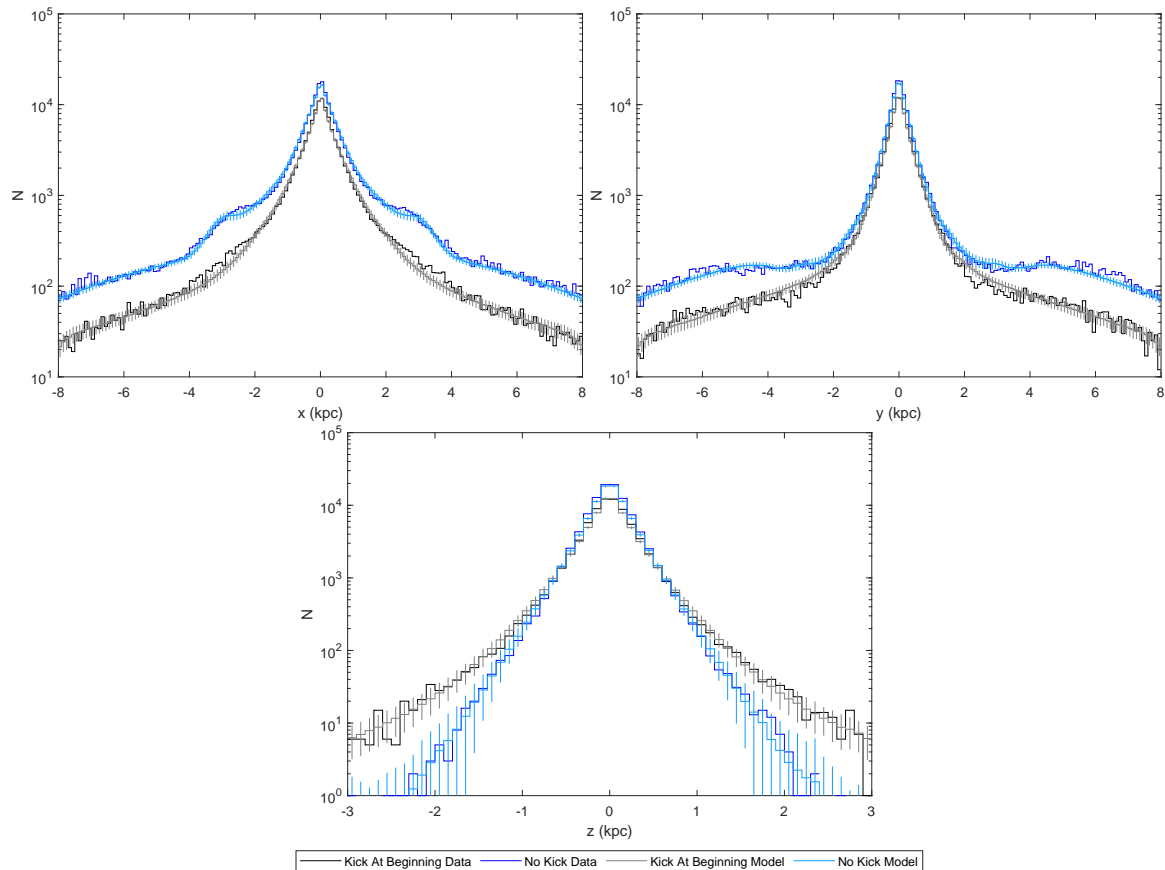


Fig. 3.5 MWa profile along  $x$ ,  $y$  and  $z$  axes with kicks occurring at the beginning. We show both  $N$ -body simulation data and data simulated using the fitted model. For the fitted model we show the mean number of particles in each bin and the standard deviation.

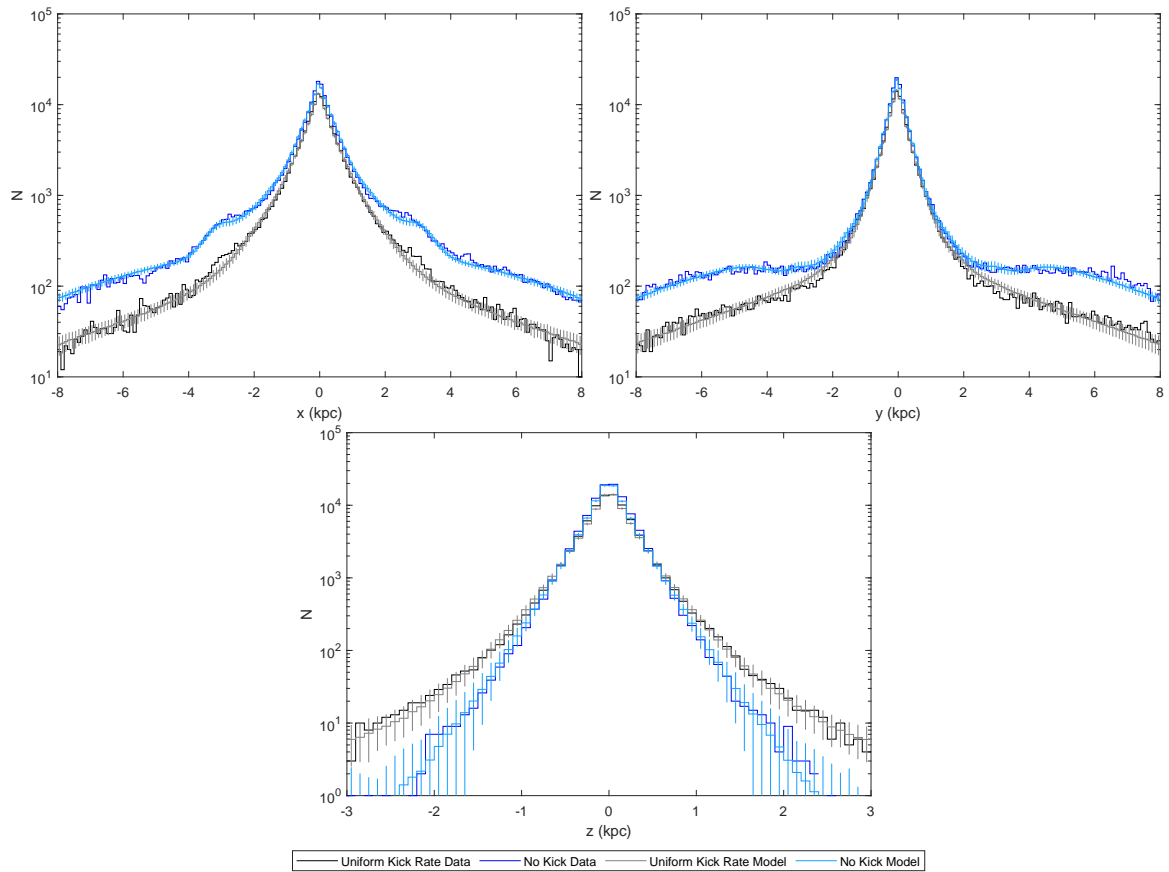


Fig. 3.6 MWa profile along  $x$ ,  $y$  and  $z$  axes with a uniform kick rate. We show both  $N$ -body simulation data and data simulated using the fitted model. For the fitted model we show the mean number of particles in each bin and the standard deviation.

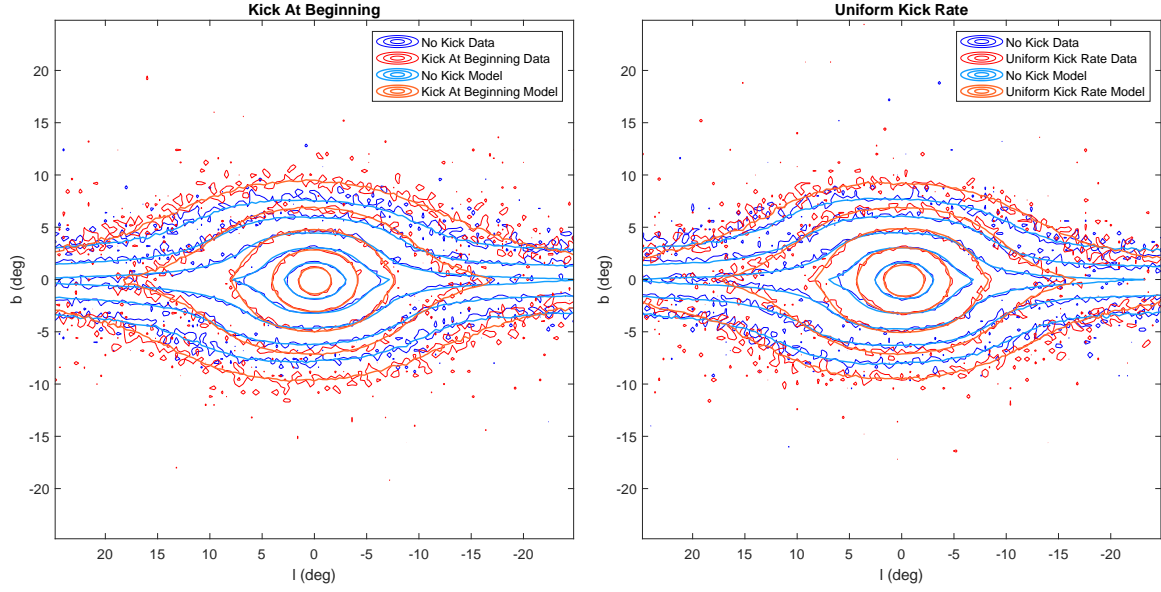


Fig. 3.7 MWA flux distribution in Galactic coordinates. The contours for each distribution are at 1, 2, 4, 8 and 16 times the mean in this region. The Sun is placed at a distance of 7.9 kpc, at an angle relative to the bar of  $20^\circ$  and at a height of 15 pc.

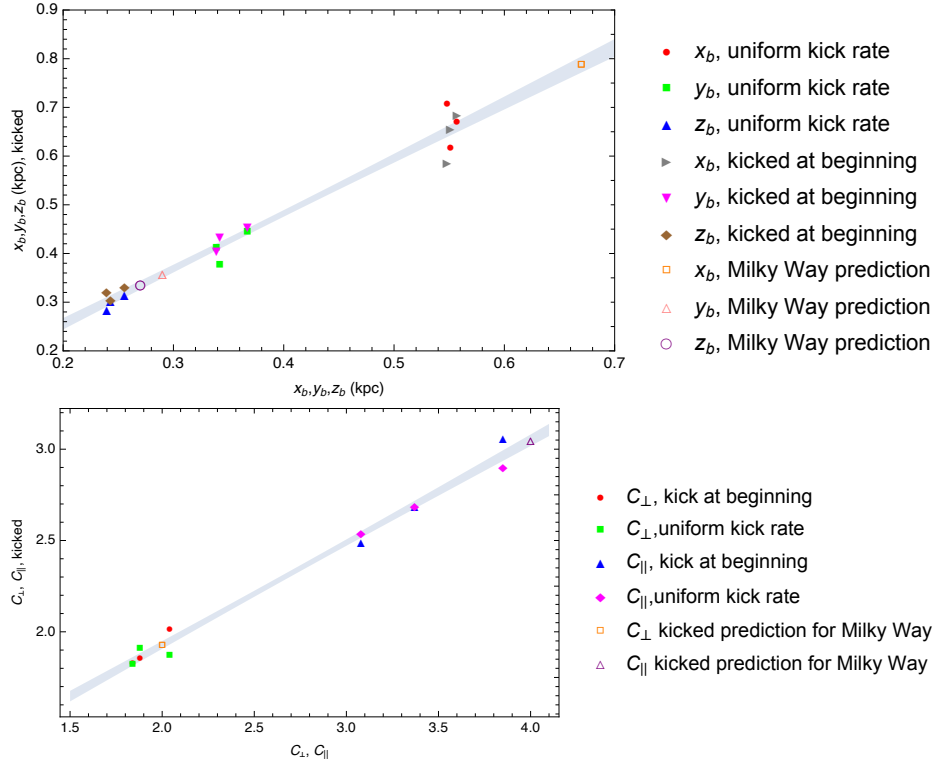


Fig. 3.8 Simulation parameters with 68% confidence interval bands for straight line model fits. The closed symbol values are obtained from values given in Tables 3.2, 3.3, and 3.4. The predictions for the ref. [26] model of the Milky Way Galaxy are given as open symbols.

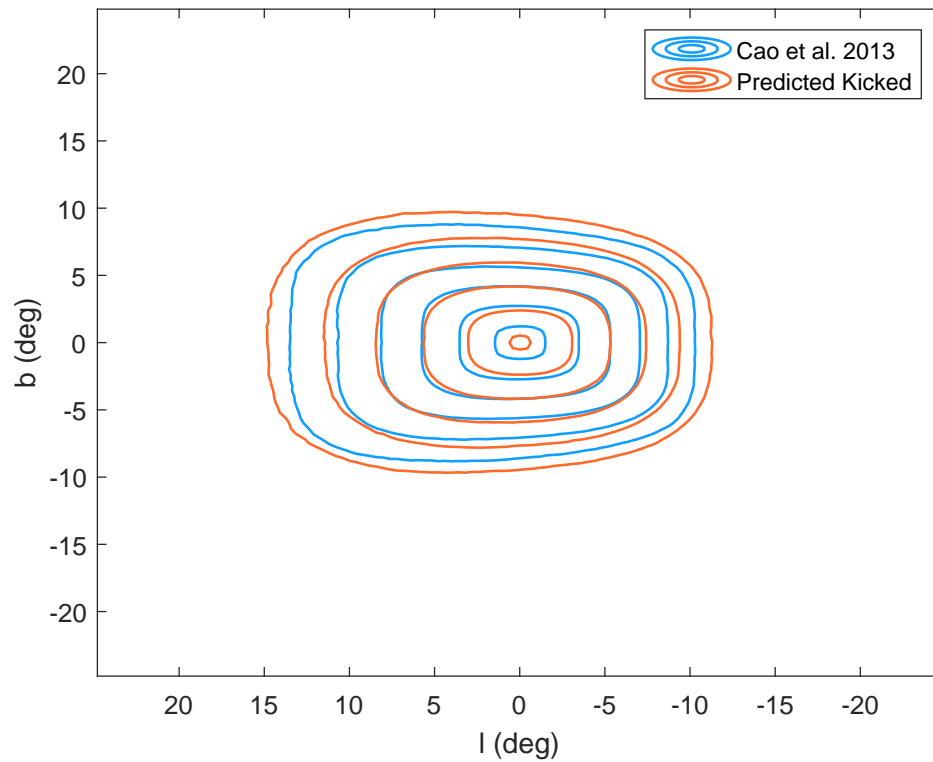


Fig. 3.9 Line of sight contours for model of the Milky Way given in ref. [26]. Both the model and its kicked version, obtained from the parameters in Table 3.7, are shown. As in ref. [26] the Galactic Centre is taken to be 8.13 kpc away and the angle of the bulge to be  $29^\circ$ . Contours are from 1 to 32 times the mean of the corresponding case given in steps of factors of 2.

then reran the simulations with a larger number of particles at the required kick velocity scale and used MCMC to fit the data with a model.

We used three models intended to approximate the Milky Way, these were the MWa, MWb and MWc0.8 models of Fujii et al. [129]. Our results were consistent with theirs as can be seen, for example, by comparing our Fig. 3.1 to the top left hand panels of their Figs. 1, 2, and 3. In Cao et al. [26] the bar scale lengths for a modified Bessel function of the second kind model fitted to red clump giant data are 0.67, 0.29 and 0.27 for the  $x$ ,  $y$  and  $z$  axes respectively, with the parameters  $C_{\parallel}$  and  $C_{\perp}$  fixed at 4 and 2. Our fits to  $N$ -body models without kicks find  $(x_b, y_b, z_b)$  of (0.56, 0.37, 0.26) for MWa, (0.55, 0.34, 0.24) for MWb and (0.55, 0.34, 0.24) for MWc0.8. We also have  $C_{\parallel}$  between 3 and 4, producing a boxy structure in  $x$ - $z$  and  $y$ - $z$ , this is visible in Fig. 3.2 for the MWa case and Figs. 3.3 and 3.4 for the MWb and MWc0.8 cases. The other shape parameter  $C_{\perp}$  was relatively close to 2 in all cases, resulting in a more elliptical shape in  $x$ - $y$ . We found it was necessary to extend the bar structure using the long bar component given in Eq. 3.11. Without it the bar scale parameters would be larger, but the bar would not be long enough to explain the structure seen for  $2 \text{ kpc} \lesssim |x| \lesssim 4 \text{ kpc}$  in Fig. 3.5 in the MWa, kicked at the beginning case, and Figs. 3.6, E.1, E.2, E.3 and E.4 in the uniform kick rate, MWb and MWc0.8 cases. We find the disk scale height to be 0.22 kpc for the MWa, MWb, and MWc0.8 unkicked cases. This is at the lower end of the range of 220 to 450 pc given in Bland-Hawthorn and Gerhard [107]. There is also a small spherically symmetric Hernquist component  $\sim 1\%$  of the particles in the region of interest for the unkicked cases.

Eckner et al. [127] argued using the virial theorem that kicks  $\langle v^2 \rangle \lesssim (70 \text{ km s}^{-1})^2$  would lead to a “smoothing” of the distribution of 700–900 pc. We show the effect of a 400 pc and an 800 pc Gaussian smoothing on the fitted bulge (bar plus Hernquist bulge) distribution in Figs. 3.10 and 3.12 for the kicked at the beginning case. The uniform kick rate case is shown in Fig. 3.11. Those figures also show the bulge component of the models with and without kicks for comparison. It is clear that a Gaussian smoothing kernel will remove the peak that survives in the  $N$ -body simulations of kicked distributions and, particularly for the 800 pc case, will produce an apparently spherically symmetric bulge. From the peculiar velocity data we inferred kicks that are larger than assumed by Eckner et al. [127] with  $\sigma_k$  at around  $80\text{--}100 \text{ km s}^{-1}$  ( $\langle v^2 \rangle = 3\sigma_k^2$  for a Maxwell distribution where the angular brackets signify the mean value) so the smoothing effect of the Gaussian would be even more severe. We show an even smaller Gaussian smoothing of 200 pc in Figs. 3.13 and 3.15 for the kicked at beginning case and in Fig. 3.14 for the uniform kick rate case. We also show in Fig. 3.16 the profile for particles with smaller kick scales between  $0 \text{ km s}^{-1}$  and  $80 \text{ km s}^{-1}$  for the kicked at the beginning case. The corresponding uniform kick rate case is shown in Fig. 3.17. In



these two figures, each kick scale has only  $4 \times 10^5$  particles; therefore, to reduce noise, the bins in each of the other two dimensions are twice as big as in previous single dimensional plots, and particles within 0.5 kpc (previously 0.25 kpc) of the axis are included. The profiles along the  $z$  axis in particular show that there is a reduction in the slope as the kick velocities increase, along the other two axes the general increase in scaleheight is seen as a reduction in density. These results demonstrates that Gaussian smoothing is not a good way of modelling a kicked version of a boxy bulge template.

In every case, the bar fitted to the kicked data is both broader, with larger scale parameters  $x_b$ ,  $y_b$  and  $z_b$ , and less boxy, with smaller  $C_{\parallel}$ . For example, for model MWa ( $x_b, y_b, z_b$ ) increases from (0.56, 0.37, 0.26) to (0.69, 0.46, 0.33) and (0.67, 0.45, 0.32) for the kick at beginning case and the uniform kick rate case respectively, while  $C_{\parallel}$  declines from 3.08 to 2.49 and 2.54. The spherically symmetric Hernquist bulge increases from  $\sim 1\%$  of the particles to 6% for MWa. Like the bar, it becomes broader with  $a_b$  increasing from around 0.2 kpc to 0.51 kpc and 0.56 kpc. For the other two models similar changes occur,  $P(\text{Long Bar})$  and  $a_b$  both increase significantly. In MSP model A1 of Chapter 2, the disk parameters were  $\sigma_r = 4.5^{+0.5}_{-0.4}$  kpc and  $z_0 = 0.71^{+0.11}_{-0.09}$ . In the current chapter, after being kicked, the disk scale heights  $z_0$  of all models increase from 0.22 kpc to  $\gtrsim 1$  kpc, while  $\sigma_r$  is in the range 5–7 kpc. However, we find that in all kicked cases, except for MWb with kicks occurring at the beginning, the long bar behaves like a relatively thin disk component. We have  $x_{\text{lb}} \approx y_{\text{lb}}$ ,  $C_{\perp, \text{lb}} \sim 2$  and  $R_{\text{out}} \gtrsim 7$ , resulting in a density  $\sim \exp(-R/R_0)$  in  $R$  for scalelength  $R_0$ . These scalelengths would then be between approximately 2.4 kpc and 2.8 kpc. For comparison, in Bland-Hawthorn and Gerhard [107] the Milky Way disk scalelength is reported as  $2.6 \pm 0.5$  kpc. The exponential scaleheights of these “long bars” range between about 0.5–0.6 kpc. In Fig. 3.5, for the MWa kicked at the beginning case and in Figs. 3.6, E.2, E.3 and E.4 for the uniform kick rate and MWb, and MWc0.8 case, there may be, to varying degrees, an excess of kicked particles over the model in the region of  $2 \text{ kpc} \lesssim |x| \lesssim 4 \text{ kpc}$ . However these are far less prominent than the long bar of the data that has not been kicked.

Our main aim in this chapter was to estimate the effect of the MSP kicks on their distribution in the Milky Way Galactic bulge. However, comparing to Cao et al. [26], none of our simulations had quite the right bulge parameters. But, there appears to be a linear relationship between the unkicked scale parameters  $x_b$ ,  $y_b$ , and  $z_b$  and their kicked counterparts. Similarly, there appears to be a linear relationship between  $C_{\perp}$ ,  $C_{\parallel}$ , and their kicked counterparts. Therefore, we were able to estimate the Milky Way Galactic bulge kicked parameters as shown in Fig. 3.8 and Table 3.7. As can be seen, there is more scatter in the  $x_b$  parameter. This is not unexpected due to the already mentioned degeneracy with the long bar. Also, as can be seen, estimating the Milky Way bulge kicked  $x_b$  parameter did

involve a reasonable amount of extrapolation and so future simulations which have a larger  $x_b$  will be needed to check it. A made-to-measure [138, 139] approach may be needed. This would also be advantageous as it could take into account the X/peanut shaped morphology of the bulge [29, 140, 141] as done in ref. [142]. This would be particularly beneficial as there is some preliminary evidence that the X-shape may improve the fit to the Fermi-LAT gamma-ray data [10].

In conclusion, we used  $N$ -body simulations to explore the effect of a Maxwell distributed kick on the distribution of pulsars in the Galactic Center. We find that while a 700–900 pc Gaussian smoothing of the stellar mass would be too aggressive, the bulge distribution of the kicked particles is slightly broader and less boxy. From these results, we expect that the GCE would not be exactly correlated with the stellar mass in the Galactic Center. As can be seen from Table 3.7, we would not expect the GCE to appear spherically symmetric due to the MSP kicks as that would require  $x_b = y_b = z_b$  and  $C_{\perp} = C_{\parallel} = 2$  which are far from our inferred points relative to their error bars.

The amount of spatial smoothing of the bulge MSPs will depend on the proportion of MSPs in the bulge that are made from the recycling channel and the proportion that are made from the accretion induced collapse channel. Motivated by similarities between the bulge and disk population seen in Chapter 2 we have assumed this mixture is the same as the disk MSPs. If the GCE is due to bulge MSPs, its morphology could be used to check our smoothing prediction by comparing if there are any deviations between the GCE morphology and the stellar spatial distribution. A complication to this approach would be the possibility of some smearing of the GCE due to cosmic ray electron diffusion [143, 144]. An additional complication is that if the MSP is spun up by a captured star then the MSP spatial distribution would be proportional to the stellar density squared [8, 127]. We have been assuming that the MSPs formed in a binary system and so have a density proportional to the stellar density. Eventually, once the bulge MSPs are resolved [109, Chapter 2], comparing their spatial distribution to the stellar distribution should provide independent information to more robustly estimate the natal kick distribution.

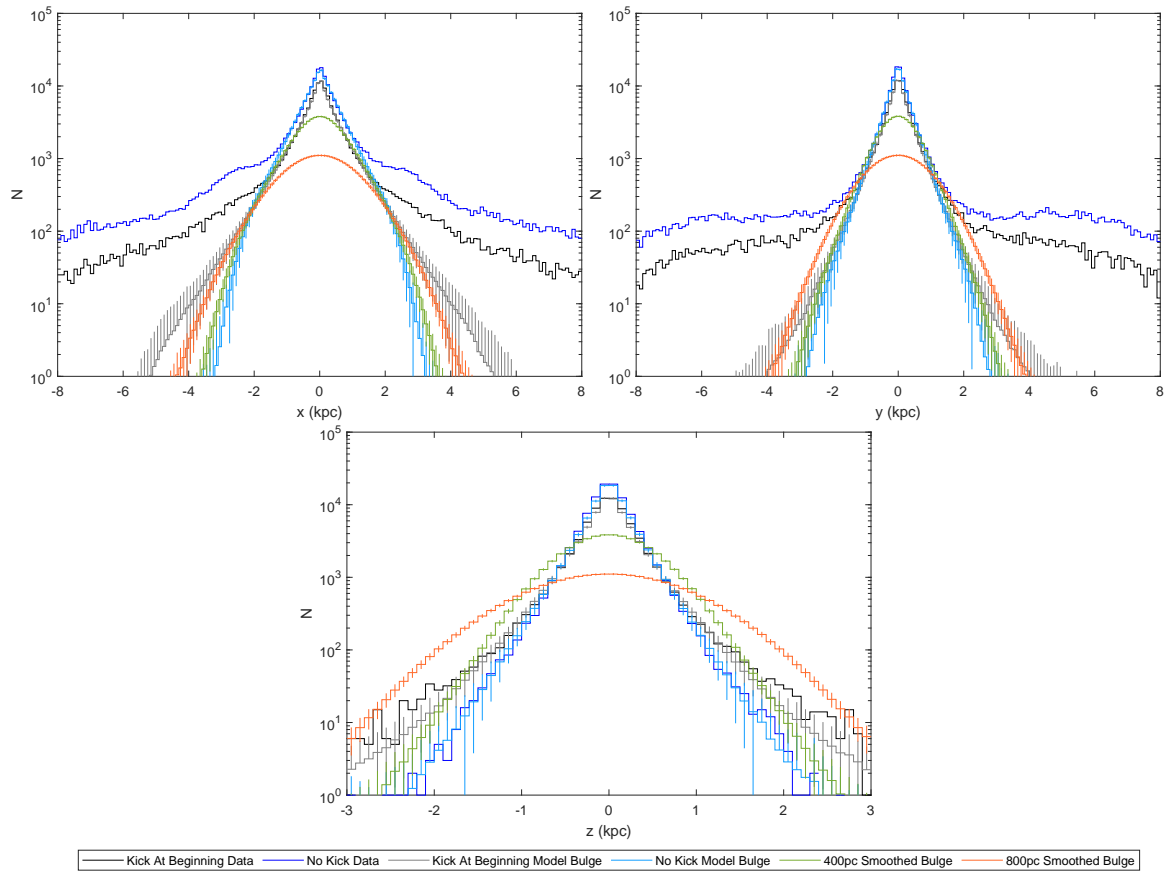


Fig. 3.10 MWa profile along  $x$ ,  $y$  and  $z$  axes with kicks occurring at the beginning. Here we show the fitted bulge components as well as the no kick bulge smoothed with 400 pc and 800 pc Gaussians. We show both  $N$ -body simulation data and data simulated using the fitted model. For the fitted model we show the mean number of particles in each bin and the standard deviation.

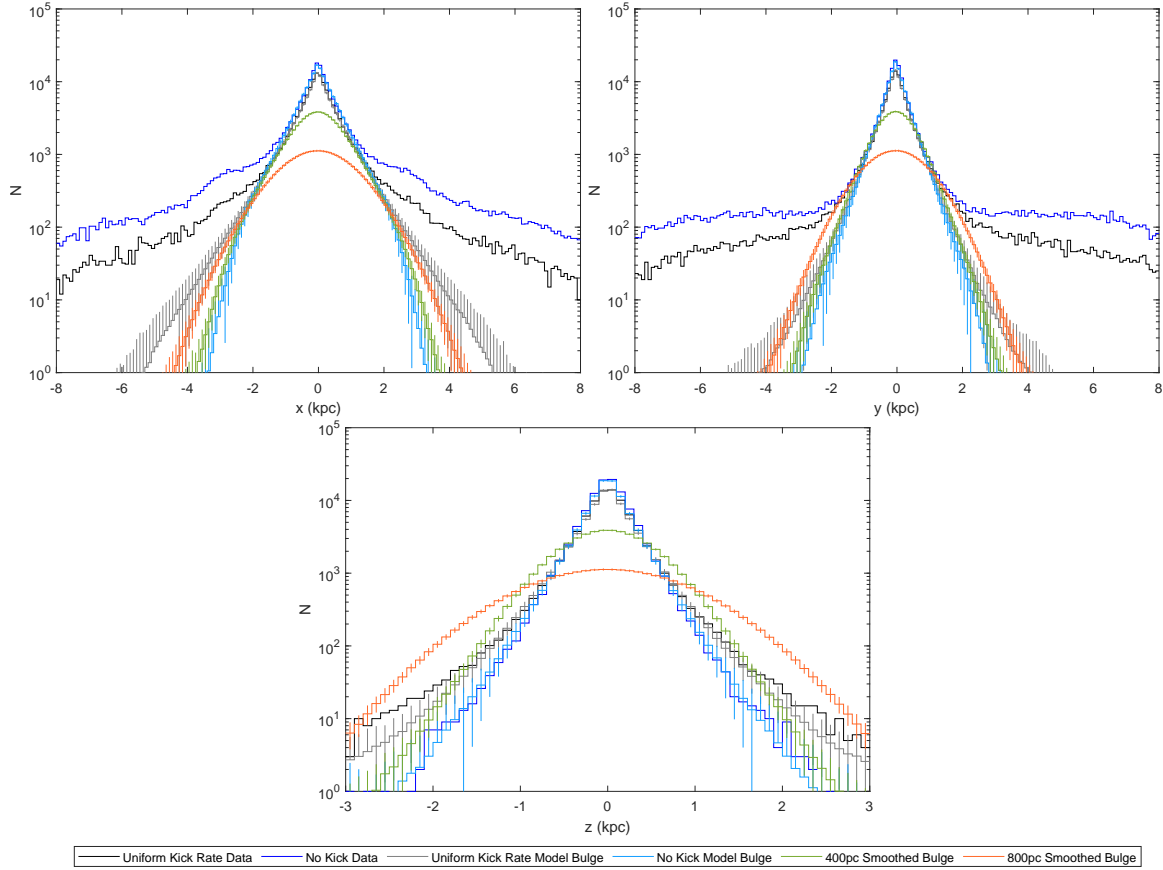


Fig. 3.11 MWa profile along  $x$ ,  $y$  and  $z$  axes with a uniform kick rate. Here we show the fitted bulge components as well as the no kick bulge smoothed with 400 pc and 800 pc Gaussians. We show both  $N$ -body simulation data and data simulated using the fitted model. For the fitted model we show the mean number of particles in each bin and the standard deviation.

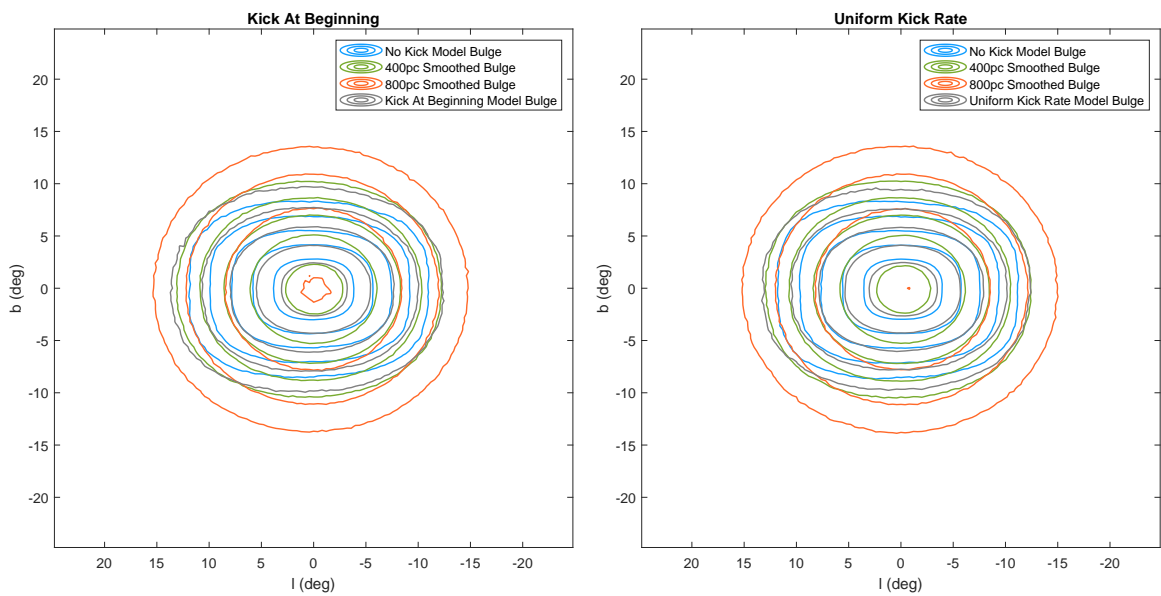


Fig. 3.12 MWA bulge flux distribution in Galactic coordinates. We also show the no kick bulge smoothed with 400 pc and 800 pc Gaussians. The contours for each distribution are at 1, 2, 4, 8 and 16 times the mean in this region. The Sun is placed at a distance of 7.9 kpc, at an angle relative to the bar of  $20^\circ$  and at a height of 15 pc.

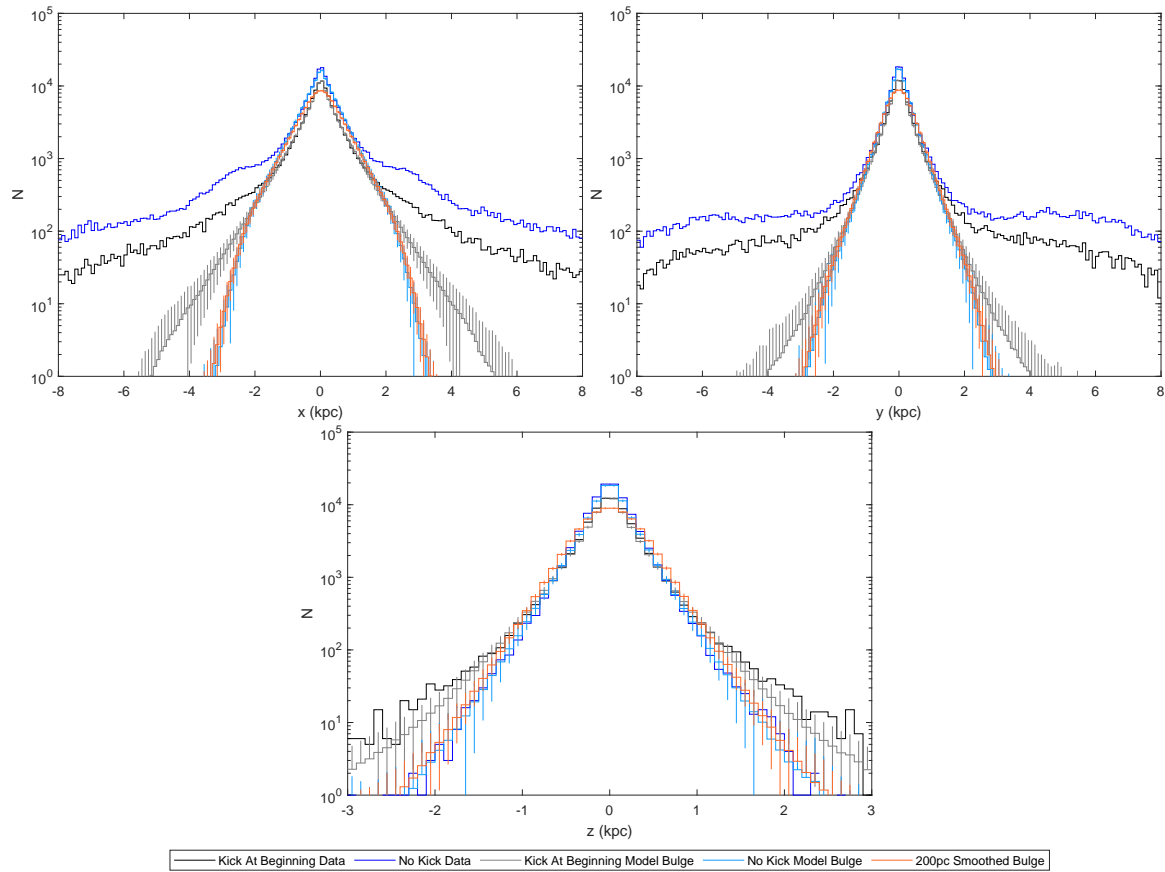


Fig. 3.13 MWa profile along  $x$ ,  $y$  and  $z$  axes with kicks occurring at the beginning. Here we show the fitted bulge components as well as the no kick bulge smoothed with a 200 pc Gaussian. We show both  $N$ -body simulation data and data simulated using the fitted model. For the fitted model we show the mean number of particles in each bin and the standard deviation.

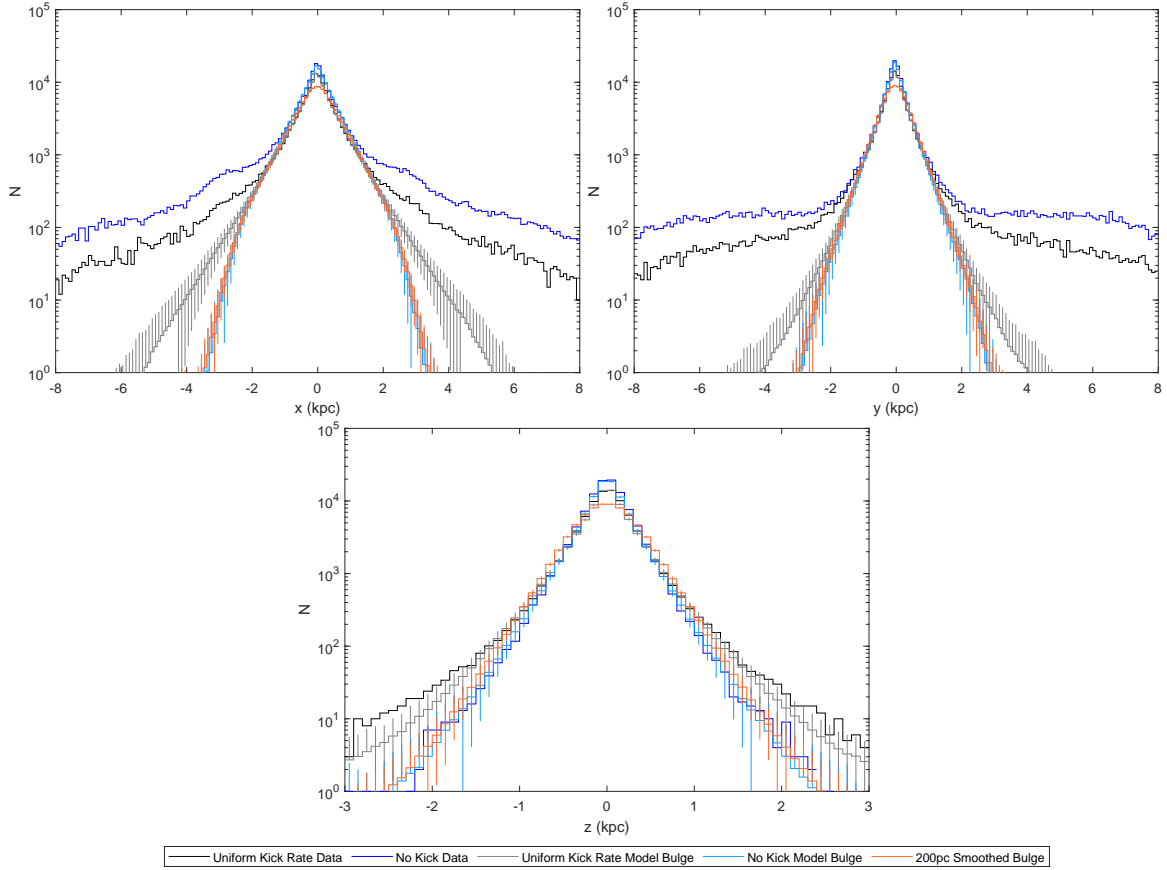


Fig. 3.14 MWa profile along  $x$ ,  $y$  and  $z$  axes with a uniform kick rate. Here we show the fitted bulge components as well as the no kick bulge smoothed with a 200 pc Gaussian. We show both  $N$ -body simulation data and data simulated using the fitted model. For the fitted model we show the mean number of particles in each bin and the standard deviation.

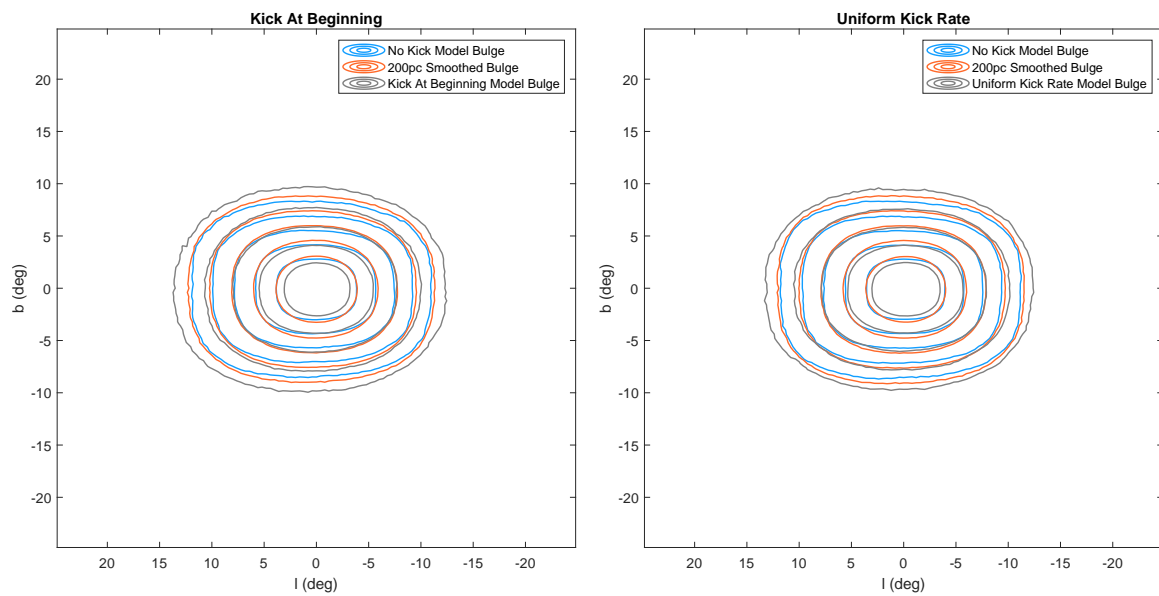


Fig. 3.15 MWA bulge flux distribution in Galactic coordinates. We also show the no kick bulge smoothed with a 200 pc Gaussian. The contours for each distribution are at 1, 2, 4, 8 and 16 times the mean in this region. The Sun is placed at a distance of 7.9 kpc, at an angle relative to the bar of  $20^\circ$  and at a height of 15 pc.



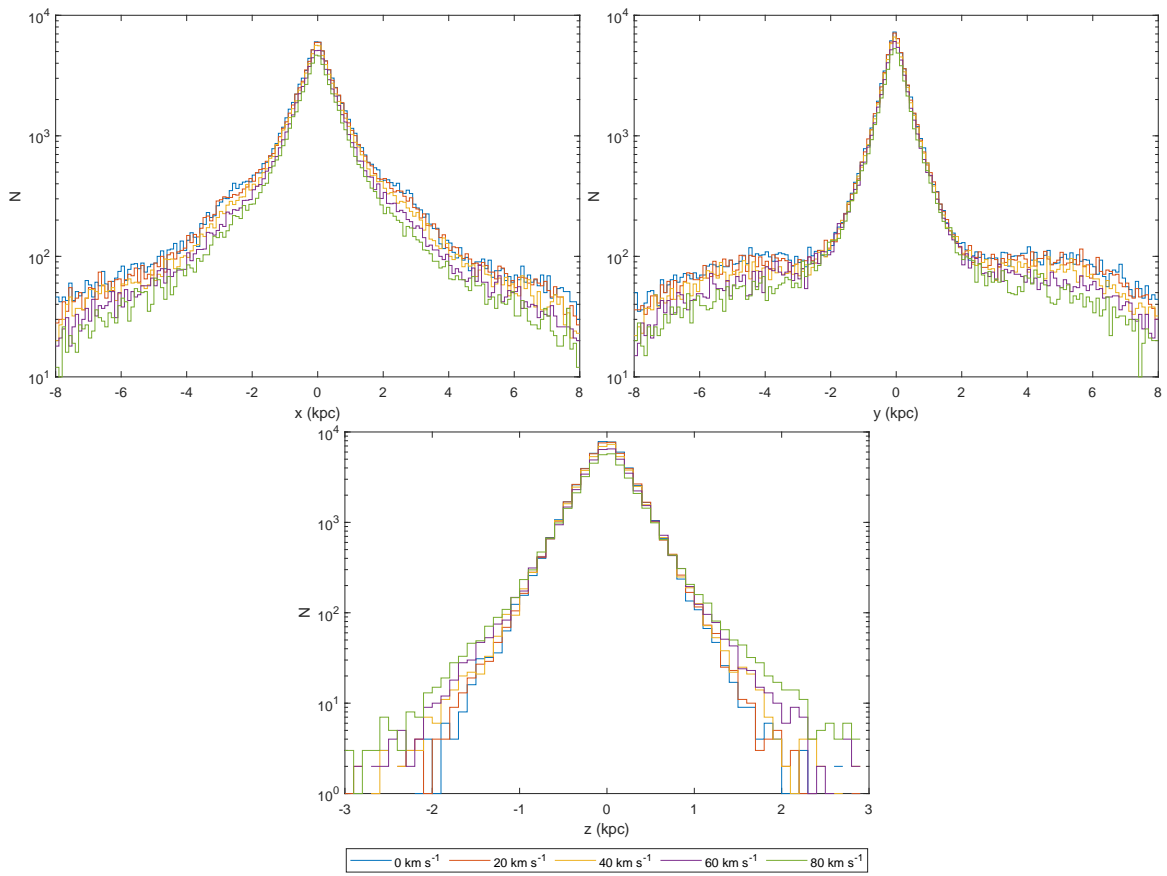


Fig. 3.16 MWa profile along  $x$ ,  $y$  and  $z$  axes with kicks occurring at the beginning for kicks between  $0 \text{ km s}^{-1}$  and  $80 \text{ km s}^{-1}$  in  $20 \text{ km s}^{-1}$  increments.

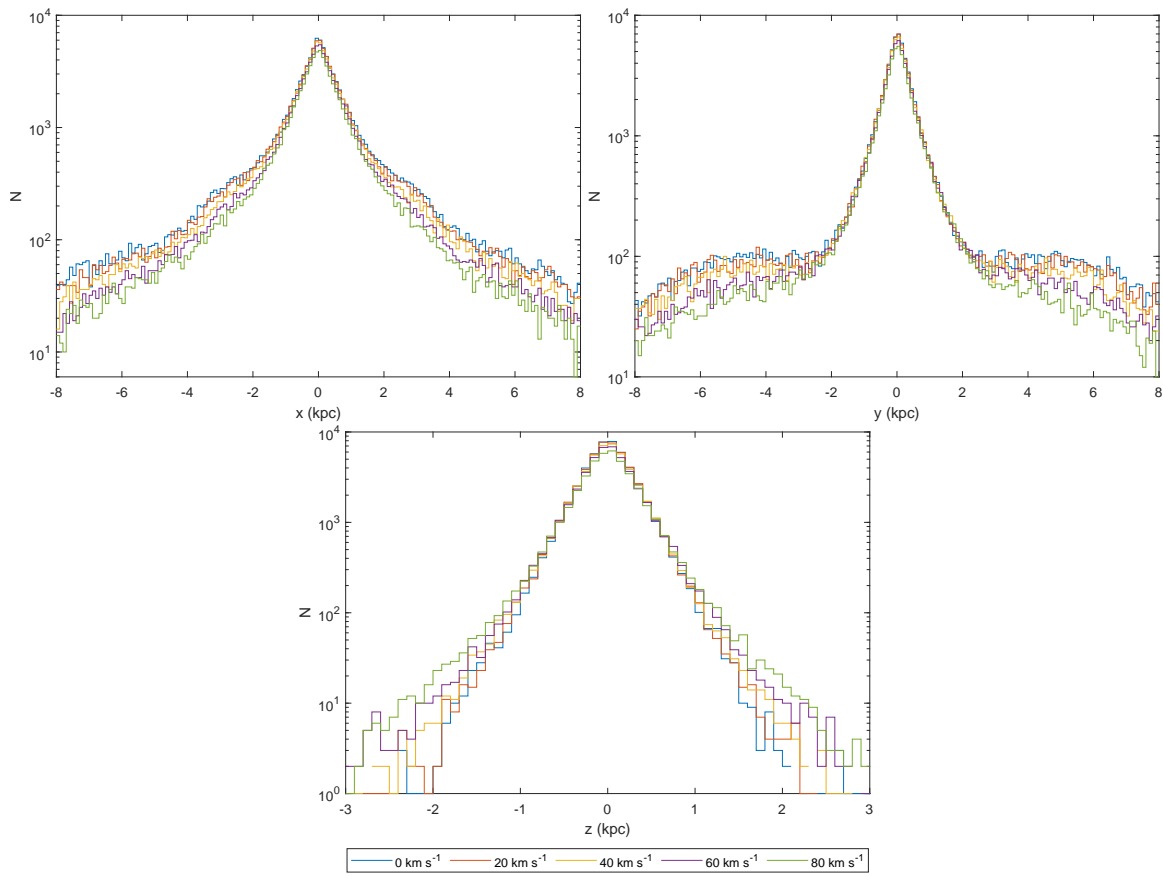


Fig. 3.17 MWa profile along  $x$ ,  $y$  and  $z$  axes with a uniform kick rate for kicks between  $0 \text{ km s}^{-1}$  and  $80 \text{ km s}^{-1}$  in  $20 \text{ km s}^{-1}$  increments.

# Chapter 4

## Summary and Future Work

In this thesis we have explored the possibility that the GCE, an extended gamma-ray source found in the Fermi-LAT data, is produced by a bulge population of unresolved MSPs. Resolved MSPs have gamma-ray spectra that peak at a few GeV, like the GCE, and they have a luminosity distribution where they would be unlikely to be resolved at the distance of the Galactic Center. Additionally, the GCE now appears to have a spatial distribution similar to that of the Galactic bulge, indicating the source of the GCE is a population of unresolved astrophysical point sources. This disfavors the self-annihilating WIMPs scenario in which the distribution would be spherically symmetric.

In Chapter 2, we used MSPs with confirmed gamma-ray pulsations detected in the Fermi-LAT data to model the Milky Way MSP population. Combining data in the Fermi Large Area Telescope fourth source catalog data release 2 [4FGL-DR2: 89] and the ATNF pulsar catalog [52], we used MCMC to fit a model to their positions in galactic longitude  $l$ , latitude  $b$ , distance  $d$ , period  $P$ , period derivative  $\dot{P}$ , proper motions  $\mu_l$  and  $\mu_b$ , gamma-ray flux  $F$ , spectral energy cut-off  $E_{\text{cut}}$ , and spectral index  $\Gamma$ . Our model consisted of a disk, boxy bulge and nuclear bulge, where the latter two components were responsible for a simulated GCE. We performed all fits both with and without the GCE.

We tried various models of the luminosity distribution. Ranking models using the WAIC, there was a clear preference for a dependence of luminosity on  $E_{\text{cut}}$ , the magnetic field strength  $B$  and the spin down power  $\dot{E}$ . For the form  $L \propto E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$ , we found  $a_\gamma = 1.2 \pm 0.3$ ,  $b_\gamma = 0.1 \pm 0.4$  and  $d_\gamma = 0.5 \pm 0.1$ . This was consistent with the prediction for curvature radiation of Kalapotharakos et al. [18] where  $a_\gamma = 4/3$ ,  $b_\gamma = 1/6$  and  $d_\gamma = 5/12$ . It is inconsistent with their prediction for synchrotron radiation of  $a_\gamma = 1$ ,  $b_\gamma = 0$  and  $d_\gamma = 1$ . The worst model, with  $\Delta\text{WAIC} = 42$  relative to the best model when the GCE was included

in the fit (and  $\Delta\text{WAIC} = 47$  otherwise), was one in which the luminosity was log-normally distributed and independent of any other property of an MSP.

Using a formula relating magnetic field strength, initial period and age to current period, we can model the present day period and period derivative distribution. We modelled MSPs as having log-normally distributed magnetic field strengths and initial periods, these distributions were fitted to the data. For our best model, we found for the magnetic field strength distribution that  $\log_{10}(B/\text{G})$  had a mean of  $8.21^{+0.05}_{-0.06}$  with a standard deviation of  $0.21^{+0.03}_{-0.02}$  where  $B$  has units of Gauss. For the initial periods with units of seconds,  $\log_{10}(P_i/\text{s})$  had a mean of  $-2.61^{+0.05}_{-0.04}$  and a standard deviation of  $0.13^{+0.02}_{-0.02}$ . Aside from their spatial distributions, the MSP populations making up each of the three components differed only in their SFR. Fitting a five bin DTD shared between the components, where the delay time is the time between star and MSP formation, we attempted to determine the age distribution of MSPs in each component. As an alternative age distribution, we had a uniform birth rate between the present and 10 Gyr ago for all three MSP populations. Comparing these cases with the WAIC, we found that the case where the age distributions of the disk, boxy bulge and nuclear bulge were different was not significantly better than the uniform birth rate case. The fitted age distribution was poorly constrained and not clearly inconsistent with a uniform birth rate.

For our best model with  $L \propto E_{\text{cut}}^{a_\gamma} B^{b_\gamma} \dot{E}^{d_\gamma}$  and with the fitted DTD, we found the GCE could be produced by a bulge population of  $N_{\text{bulge}} = 34^{+17}_{-10}$  thousand MSPs. The probability that none of these bulge MSPs have been resolved at present was 0.16 with a 68% interval of 1–4. We identified three resolved MSPs with probability greater than 0.05 of being bulge MSPs: PSR J1747-4036 with probability 0.4, PSR J1811-2405 with probability 0.5 and PSR J1855-1436 with probability 0.1. However, these probabilities require that the distances of these pulsars really are large enough to locate them in a region of space where the bulge density is significant. That density is determined by the bulge spatial model and by the assumption that MSPs are the source of the GCE. Switching to a  $r^{-2.4}$  bulge profile with a cutoff at  $r = 3.1$  kpc changed the previously mentioned probabilities to 0.5, 0.2, and  $4 \times 10^{-4}$  respectively.

In future, as the number of gamma-ray MSPs detected by Fermi-LAT increases over time, the constraints on the model parameters may be significantly improved. However, our model of the Galactic MSP population could also be improved. We included a randomly distributed magnetic axis angle  $\alpha$  relative to the rotation axis as this affects the rate at which a pulsar spins down for a given magnetic field strength. Although we have a parameter  $\eta$  that allows for random variation in pulsar luminosity, we do not explicitly account for the possibility that the flux may vary due to  $\alpha$  as well as the viewing angle relative to rotation

axis. This may mean the resolved MSPs are biased towards particular values of  $\alpha$ . This could be accounted for by including a model of the gamma-ray emission geometry. Another improvement would be to include radio detected MSPs in the fit. This would increase the number of MSPs involved in constraining the model parameters. They would require a model of the radio luminosity and of the radio flux threshold. Almost all of the gamma-ray MSPs we have used in our analysis were discovered in radio before they were found in the Fermi-LAT data. As many may have been detected by Fermi-LAT only because of that previous discovery, including a radio model may allow explicit inclusion of a radio detection threshold component to the gamma-ray flux detection threshold. Currently, to model the possibility that MSPs with gamma-ray flux near the threshold may or may not be resolved for this reason, we simply allow the galactic longitude and latitude dependent flux threshold to vary log-normally, finding  $\sigma_{\text{th}} = 0.28^{+0.05}_{-0.04}$ . However, it may be that the radio threshold depends very differently on position in the sky.

As part of the fitted model in Chapter 2 we included a simple model of pulsar velocities in which they were travelling on circular orbits plus a Maxwell distributed peculiar velocity in a random direction. We found for the scale parameter of this peculiar velocity  $\sigma_v = 77 \pm 6$  km s<sup>-1</sup>. In Chapter 3 we investigated the effect of pulsar kicks on the structure of the bulge. We did this by running  $N$ -body simulations of a galaxy intended be similar to the Milky Way, adding in particles which were given a kick of a magnitude selected to produce a peculiar velocity distribution close to that of the resolved MSPs.

We ran  $N$ -body simulations using three different sets of initial conditions, and with kicks either occurring all at the beginning or at a uniform rate during the 10 Gyr simulation. We found initial Maxwell distributed kicks of around 80–100 km s<sup>-1</sup> depending on the initial conditions and kick rate. We fitted a parametric model of the Galaxy to the particle distributions. This model consisted of a disk, spherically symmetric bulge, long bar and a bar of the same form as in Cao et al. [26]. Eckner et al. [127] argued using the virial theorem that MSP kicks where  $\langle v^2 \rangle \lesssim (70 \text{ km s}^{-1})^2$  would lead to a smoothing scale of the spatial distribution of 700–900 pc. This would make the bulge distribution relatively spherical. We found the kicks broadened the bar structure and made it less boxy, but they did not render it spherically symmetric. The kicks did increase the spherically symmetric component at the galactic center from around  $\sim 1\%$  to 5–8% of the model, however the bar was the more significant structure in that region. Taking the fitted central values for the bar scale parameters  $x_b$ ,  $y_b$ , and  $z_b$  and separately the shape parameters  $C_{\perp}$  and  $C_{\parallel}$ , we found they lie nearly on a line when plotting the values for the case with no kick against the values for the kicked particle distribution. Using a linear fit we showed the kicked scale parameters are increased by around 10%. From this we suggested that for the Cao et

al. [26] model of the Milky Way boxy bulge fitted to the red clump giant distribution, the parameters  $(x_b, y_b, z_b, C_\perp, C_\parallel) = (0.67, 0.29, 0.27, 2, 4)$  could, for a kicked version, become  $(0.79, 0.36, 0.35, 1.93, 3.05)$ . We expect that the GCE, if it is produced by a population of unresolved MSPs, will differ from the stellar mass distribution in the Galactic Center. However, pulsar birth kicks should not result in a spherically symmetric distribution of bulge MSPs. Assuming a similar mixture of MSP formation channels applies to the bulge as the disk, so that the natal kick distributions in velocity and in time are not significantly different, future studies into the morphology of the GCE may find templates produced with a slightly larger and less boxy bar provide a better fit to the data.

# References

- [1] H. T. Freudenreich, *A COBE Model of the Galactic Bar and Disk*, *ApJ* **492** (Jan., 1998) 495–510, [astro-ph/9707340].
- [2] L. Goodenough and D. Hooper, *Possible Evidence For Dark Matter Annihilation In The Inner Milky Way From The Fermi Gamma Ray Space Telescope*, *arXiv e-prints* (Oct., 2009) arXiv:0910.2998, [0910.2998].
- [3] D. Hooper and T. Linden, *Origin of the gamma rays from the galactic center*, *Physical Review D* **84** (Dec, 2011) .
- [4] K. N. Abazajian and M. Kaplinghat, *Detection of a Gamma-Ray Source in the Galactic Center Consistent with Extended Emission from Dark Matter Annihilation and Concentrated Astrophysical Emission*, *Phys. Rev. D* **86** (2012) 083511, [1207.6047].
- [5] C. Gordon and O. Macias, *Dark Matter and Pulsar Model Constraints from Galactic Center Fermi-LAT Gamma Ray Observations*, *Phys. Rev. D* **88** (2013) 083521, [1306.5725].
- [6] O. Macias, C. Gordon, R. M. Crocker, B. Coleman, D. Paterson, S. Horiuchi et al., *Galactic bulge preferred over dark matter for the galactic centre gamma-ray excess*, *Nature Astronomy* **2** (Mar, 2018) 387–392.
- [7] R. Bartels, E. Storm, C. Weniger and F. Calore, *The Fermi-LAT GeV excess as a tracer of stellar mass in the Galactic bulge*, *Nature Astronomy* (Aug., 2018) , [1711.04778].
- [8] O. Macias, S. Horiuchi, M. Kaplinghat, C. Gordon, R. M. Crocker and D. M. Nataf, *Strong evidence that the galactic bulge is shining in gamma rays*, *JCAP* **2019** (Sept., 2019) 042, [1901.03822].
- [9] K. N. Abazajian, S. Horiuchi, M. Kaplinghat, R. E. Keeley and O. Macias, *Strong constraints on thermal relic dark matter from Fermi-LAT observations of the Galactic Center*, *Phys. Rev. D* **102** (2020) 043012, [2003.10416].

- [10] B. Coleman, D. Paterson, C. Gordon, O. Macias and H. Ploeg, *Maximum entropy estimation of the Galactic bulge morphology via the VVV Red Clump*, *MNRAS* **495** (May, 2020) 3350–3372, [1911.04714].
- [11] M. Di Mauro, *Characteristics of the Galactic Center excess measured with 11 years of Fermi-LAT data*, *Phys. Rev. D* **103** (Mar., 2021) 063029, [2101.04694].
- [12] K. N. Abazajian, *The Consistency of Fermi-LAT Observations of the Galactic Center with a Millisecond Pulsar Population in the Central Stellar Cluster*, *JCAP* **1103** (2011) 010, [1011.4275].
- [13] D. Hooper and G. Mohlabeng, *The Gamma-Ray Luminosity Function of Millisecond Pulsars and Implications for the GeV Excess*, *JCAP* **1603** (2016) 049, [1512.04966].
- [14] D. Haggard, C. Heinke, D. Hooper and T. Linden, *Low mass x-ray binaries in the inner galaxy: implications for millisecond pulsars and the gev excess*, *Journal of Cosmology and Astroparticle Physics* **2017** (May, 2017) 056–056.
- [15] H. Ploeg, C. Gordon, R. Crocker and O. Macias, *Consistency Between the Luminosity Function of Resolved Millisecond Pulsars and the Galactic Center Excess*, *JCAP* **1708** (2017) 015, [1705.00806].
- [16] R. T. Bartels, T. D. P. Edwards and C. Weniger, *Bayesian model comparison and analysis of the Galactic disc population of gamma-ray millisecond pulsars*, *MNRAS* **481** (Dec., 2018) 3966–3987, [1805.11097].
- [17] H. Ploeg, C. Gordon, R. Crocker and O. Macias, *Comparing the galactic bulge and galactic disk millisecond pulsars*, *JCAP* **2020** (Dec., 2020) 035, [2008.10821].
- [18] C. Kalapotharakos, A. K. Harding, D. Kazanas and Z. Wadiasingh, *A fundamental plane for gamma-ray pulsars*, *The Astrophysical Journal* **883** (Sep, 2019) L4.
- [19] P. L. Gonthier, A. K. Harding, E. C. Ferrara, S. E. Frederick, V. E. Mohr and Y.-M. Koh, *Population Syntheses of Millisecond Pulsars from the Galactic Disk and Bulge*, *ApJ* **863** (Aug., 2018) 199, [1806.11215].
- [20] H. Ploeg and C. Gordon, *The effect of kick velocities on the spatial distribution of millisecond pulsars and implications for the galactic center excess*, *arXiv e-prints* (May, 2021) , [2105.13034].
- [21] J. L. Weiland, R. G. Arendt, G. B. Berriman, E. Dwek, H. T. Freudenreich, M. G. Hauser et al., *COBE Diffuse Background Experiment Observations of the Galactic Bulge*, *ApJ* **425** (Apr., 1994) L81.



- [22] E. Dwek, R. G. Arendt, M. G. Hauser, T. Kelsall, C. M. Lisse, S. H. Moseley et al., *Morphology, near-infrared luminosity, and mass of the Galactic bulge from COBE DIRBE observations*, *ApJ* **445** (June, 1995) 716–730.
- [23] J. Binney, O. Gerhard and D. Spergel, *The photometric structure of the inner Galaxy*, *MNRAS* **288** (June, 1997) 365–374, [astro-ph/9609066].
- [24] K. Z. Stanek, A. Udalski, M. Szymanski, J. Kaluzny, M. Kubiak, M. Mateo et al., *Modeling the galactic bar using red clump giants*, *Astrophys. J.* **477** (1997) 163–175, [astro-ph/9605162].
- [25] N. J. Rattenbury, S. Mao, T. Sumi and M. C. Smith, *Modelling the Galactic bar using OGLE-II red clump giant stars*, *MNRAS* **378** (July, 2007) 1064–1078, [0704.1614].
- [26] L. Cao, S. Mao, D. Nataf, N. J. Rattenbury and A. Gould, *A New Photometric Model of the Galactic Bar using Red Clump Giants*, *Mon. Not. Roy. Astron. Soc.* **434** (2013) 595–605, [1303.6430].
- [27] C. Wegg and O. Gerhard, *Mapping the three-dimensional density of the Galactic bulge with VVV red clump stars*, *MNRAS* **435** (Nov., 2013) 1874–1887, [1308.0593].
- [28] I. T. Simion, V. Belokurov, M. Irwin, S. E. Koposov, C. Gonzalez-Fernandez, A. C. Robin et al., *A parametric description of the 3D structure of the Galactic bar/bulge using the VVV survey*, *MNRAS* **471** (Nov., 2017) 4323–4344, [1707.06660].
- [29] C. Wegg, O. Gerhard and M. Portail, *The structure of the Milky Way’s bar outside the bulge*, *MNRAS* **450** (July, 2015) 4050–4069, [1504.01401].
- [30] R. Launhardt, R. Zylka and P. G. Mezger, *The nuclear bulge of the Galaxy. III. Large-scale physical characteristics of stars and interstellar matter*, *A&A* **384** (Mar., 2002) 112–139, [astro-ph/0201294].
- [31] W. B. Atwood, A. A. Abdo, M. Ackermann, W. Althouse, B. Anderson, M. Axelsson et al., *The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission*, *ApJ* **697** (June, 2009) 1071–1102, [0902.1089].
- [32] M. Ackermann, M. Ajello, A. Albert, A. Allafort, W. B. Atwood, M. Axelsson et al., *The Fermi Large Area Telescope on Orbit: Event Classification, Instrument Response Functions, and Calibration*, *ApJS* **203** (Nov., 2012) 4, [1206.1896].
- [33] O. Macias and C. Gordon, *Contribution of cosmic rays interacting with molecular clouds to the Galactic Center gamma-ray excess*, *Phys. Rev.* **D89** (2014) 063515, [1312.6671].

- [34] K. N. Abazajian, N. Canac, S. Horiuchi and M. Kaplinghat, *Astrophysical and Dark Matter Interpretations of Extended Gamma-Ray Emission from the Galactic Center*, *Phys. Rev. D* **90** (2014) 023526, [1402.4090].
- [35] T. Daylan, D. P. Finkbeiner, D. Hooper, T. Linden, S. K. N. Portillo, N. L. Rodd et al., *The characterization of the gamma-ray signal from the central Milky Way: A case for annihilating dark matter*, *Phys. Dark Univ.* **12** (2016) 1–23, [1402.6703].
- [36] FERMI-LAT collaboration, M. Ajello et al., *Fermi-LAT Observations of High-Energy  $\gamma$ -Ray Emission Toward the Galactic Center*, *Astrophys. J.* **819** (2016) 44, [1511.02938].
- [37] J. F. Navarro, C. S. Frenk and S. D. M. White, *The Structure of Cold Dark Matter Halos*, *ApJ* **462** (May, 1996) 563, [astro-ph/9508025].
- [38] E. A. Baltz, B. Berenji, G. Bertone, L. Bergström, E. Bloom, T. Bringmann et al., *Pre-launch estimates for GLAST sensitivity to dark matter annihilation signals*, *JCAP* **2008** (July, 2008) 013, [0806.2911].
- [39] D. Hooper and L. Goodenough, *Dark Matter Annihilation in The Galactic Center As Seen by the Fermi Gamma Ray Space Telescope*, *Phys. Lett. B* **697** (2011) 412–428, [1010.2752].
- [40] G. Steigman, B. Dasgupta and J. F. Beacom, *Precise relic wimp abundance and its impact on searches for dark matter annihilation*, *Physical Review D* **86** (Jul, 2012) .
- [41] W. Wang, Z. J. Jiang and K. S. Cheng, *Contribution to diffuse gamma-rays in the Galactic Centre region from unresolved millisecond pulsars*, *MNRAS* **358** (Mar., 2005) 263–269, [astro-ph/0501245].
- [42] T. D. Brandt and B. Kocsis, *Disrupted Globular Clusters Can Explain the Galactic Center Gamma Ray Excess*, *Astrophys. J.* **812** (2015) 15, [1507.05616].
- [43] O. Y. Gnedin, J. P. Ostriker and S. Tremaine, *Co-Evolution of Galactic Nuclei and Globular Cluster Systems*, *Astrophys. J.* **785** (2014) 71, [1308.0021].
- [44] Q. Yuan and B. Zhang, *Millisecond pulsar interpretation of the Galactic center gamma-ray excess*, *Journal of High Energy Astrophysics* **3** (Sept., 2014) 1–8, [1404.2318].
- [45] FERMI-LAT collaboration, A. A. Abdo et al., *The Second Fermi Large Area Telescope Catalog of Gamma-ray Pulsars*, *Astrophys. J. Suppl.* **208** (2013) 17, [1305.4385].

- [46] J. Petrović, P. D. Serpico and G. Zaharijas, *Millisecond pulsars and the Galactic Center gamma-ray excess: the importance of luminosity function and secondary emission*, *JCAP* **2015** (Feb., 2015) 023, [1411.2980].
- [47] FERMI-LAT collaboration, F. Acero et al., *Fermi Large Area Telescope Third Source Catalog*, *Astrophys. J. Suppl.* **218** (2015) 23, [1501.02003].
- [48] S. K. Lee, M. Lisanti, B. R. Safdi, T. R. Slatyer and W. Xue, *Evidence for Unresolved  $\gamma$ -Ray Point Sources in the Inner Galaxy*, *Phys. Rev. Lett.* **116** (2016) 051103, [1506.05124].
- [49] R. Bartels, S. Krishnamurthy and C. Weniger, *Strong support for the millisecond pulsar origin of the Galactic center GeV excess*, *Phys. Rev. Lett.* **116** (2016) 051102, [1506.05104].
- [50] M. Buschmann, N. L. Rodd, B. R. Safdi, L. J. Chang, S. Mishra-Sharma, M. Lisanti et al., *Foreground mismodeling and the point source explanation of the Fermi Galactic Center excess*, *Physical Review D* **102** (July, 2020) 023023, [2002.12373].
- [51] F. Calore, F. Donato and S. Manconi, *Dissecting the inner Galaxy with  $\gamma$ -ray pixel count statistics*, *arXiv e-prints* (Feb., 2021) arXiv:2102.12497, [2102.12497].
- [52] R. N. Manchester, G. B. Hobbs, A. Teoh and M. Hobbs, *The Australia Telescope National Facility pulsar catalogue*, *Astron. J.* **129** (2005) 1993, [astro-ph/0412641].
- [53] I. S. Shklovskii, *Possible Causes of the Secular Increase in Pulsar Periods.*, *Soviet Astronomy* **13** (Feb., 1970) 562.
- [54] L. Ferrario and D. Wickramasinghe, *The birth properties of Galactic millisecond radio pulsars*, *MNRAS* **375** (Mar., 2007) 1009–1016, [astro-ph/0701444].
- [55] R. M. O’Leary, M. D. Kistler, M. Kerr and J. Dexter, *Young Pulsars and the Galactic Center GeV Gamma-ray Excess*, 1504.02477.
- [56] D. Bhattacharya and E. P. J. van den Heuvel, *Formation and evolution of binary and millisecond radio pulsars*, *Physics Reports* **203** (Jan., 1991) 1–124.
- [57] D. R. Lorimer, *Binary and Millisecond Pulsars*, *Living Reviews in Relativity* **8** (Nov., 2005) 7, [astro-ph/0511258].
- [58] T. M. Tauris, N. Langer and M. Kramer, *Formation of millisecond pulsars with CO white dwarf companions - II. Accretion, spin-up, true ages and comparison to MSPs with He white dwarf companions*, *MNRAS* **425** (Sept., 2012) 1601–1627, [1206.1862].

- [59] U. Geppert and V. Urpin, *Accretion-driven magnetic field decay in neutron stars*, *MNRAS* **271** (Nov., 1994) 490.
- [60] S. Konar and D. Bhattacharya, *Magnetic field evolution of accreting neutron stars*, *MNRAS* **284** (Jan., 1997) 311–317.
- [61] I. Cholis, D. Hooper and T. Linden, *Challenges in Explaining the Galactic Center Gamma-Ray Excess with Millisecond Pulsars*, *JCAP* **1506** (2015) 043, [1407.5625].
- [62] N. Ivanova, C. O. Heinke, F. A. Rasio, K. Belczynski and J. M. Fregeau, *Formation and evolution of compact binaries in globular clusters - II. Binaries with neutron stars*, *MNRAS* **386** (May, 2008) 553–576, [0706.4096].
- [63] R. Bartels, F. Calore, E. Storm and C. Weniger, *Galactic binaries can explain the Fermi Galactic centre excess and 511 keV emission*, *MNRAS* **480** (Nov., 2018) 3826–3841, [1803.04370].
- [64] L. M. van Haaften, G. Nelemans, R. Voss, M. V. van der Sluys and S. Toonen, *Population synthesis of classical low-mass x-ray binaries in the galactic bulge*, *Astronomy & Astrophysics* **579** (Jun, 2015) A33.
- [65] T. M. Tauris, D. Sanyal, S. C. Yoon and N. Langer, *Evolution towards and beyond accretion-induced collapse of massive white dwarfs and formation of millisecond pulsars*, *A&A* **558** (Oct., 2013) A39, [1308.4887].
- [66] Wongwathanarat, A., Janka, H.-Th. and Müller, E., *Three-dimensional neutrino-driven supernovae: Neutron star kicks, spins, and asymmetric ejection of nucleosynthesis products*, *Astron. and Astrophys.* **552** (2013) A126.
- [67] E. Bear and N. Soker, *Neutron star natal kick and jets in core collapse supernovae*, *Astrophys. J.* **855** (mar, 2018) 82.
- [68] G. Hobbs, D. R. Lorimer, A. G. Lyne and M. Kramer, *A statistical study of 233 pulsar proper motions*, *Monthly Notices of the Royal Astronomical Society* **360** (Jul, 2005) 974–992.
- [69] J. Pétri, *Pulsar gamma-ray emission in the radiation reaction regime*, *MNRAS* **484** (Apr., 2019) 5669–5691, [1901.11439].
- [70] A. Gruzinov, *Aristotelian Electrodynamics solves the Pulsar: Lower Efficiency of Strong Pulsars*, *arXiv e-prints* (Mar., 2013) arXiv:1303.4094, [1303.4094].

- [71] A. K. Harding, I. A. Grenier and P. L. Gonthier, *The Geminga fraction*, *Astrophysics and Space Science* **309** (June, 2007) 221–230, [astro-ph/0703019].
- [72] C. Venter, A. K. Harding and L. Guillemot, *Probing Millisecond Pulsar Emission Geometry Using Light Curves from the Fermi/Large Area Telescope*, *ApJ* **707** (Dec., 2009) 800–822, [0911.0872].
- [73] C. Venter, T. J. Johnson and A. K. Harding, *Modeling Phase-aligned Gamma-Ray and Radio Millisecond Pulsar Light Curves*, *ApJ* **744** (Jan., 2012) 34.
- [74] M. Pierbattista, I. A. Grenier, A. K. Harding and P. L. Gonthier, *Constraining  $\gamma$ -ray pulsar gap models with a simulated pulsar population*, *A&A* **545** (Sept., 2012) A42, [1206.5634].
- [75] J. Pétri, *Theory of pulsar magnetosphere and wind*, *Journal of Plasma Physics* **82** (Oct., 2016) 635820502, [1608.04895].
- [76] P. A. Sturrock, *A Model of Pulsars*, *ApJ* **164** (Mar., 1971) 529.
- [77] M. A. Ruderman and P. G. Sutherland, *Theory of pulsars: polar gaps, sparks, and coherent microwave radiation.*, *ApJ* **196** (Feb., 1975) 51–72.
- [78] A. K. Harding, E. Tademaru and L. W. Esposito, *A curvature - radiation - pair-production model for gamma -ray pulsars.*, *ApJ* **225** (Oct., 1978) 226–236.
- [79] J. K. Daugherty and A. K. Harding, *Electromagnetic cascades in pulsars.*, *ApJ* **252** (Jan., 1982) 337–347.
- [80] J. Arons, *Pair creation above pulsar polar caps : geometrical structure and energetics of slot gaps.*, *ApJ* **266** (Mar., 1983) 215–241.
- [81] A. G. Muslimov and A. K. Harding, *Extended Acceleration in Slot Gaps and Pulsar High-Energy Emission*, *ApJ* **588** (May, 2003) 430–440, [astro-ph/0301023].
- [82] A. G. Muslimov and A. K. Harding, *High-Altitude Particle Acceleration and Radiation in Pulsar Slot Gaps*, *ApJ* **606** (May, 2004) 1143–1153, [astro-ph/0402462].
- [83] K. S. Cheng, C. Ho and M. Ruderman, *Energetic Radiation from Rapidly Spinning Pulsars. I. Outer Magnetosphere Gaps*, *ApJ* **300** (Jan., 1986) 500.
- [84] J. Chiang and R. W. Romani, *Gamma Radiation from Pulsar Magnetospheric Gaps*, *ApJ* **400** (Dec., 1992) 629.

- [85] R. W. Romani and I. A. Yadigaroglu, *Gamma-Ray Pulsars: Emission Zones and Viewing Geometries*, *ApJ* **438** (Jan., 1995) 314, [astro-ph/9401045].
- [86] R. W. Romani, *Gamma-Ray Pulsars: Radiation Processes in the Outer Magnetosphere*, *ApJ* **470** (Oct., 1996) 469.
- [87] C. Kalapotharakos, G. Brambilla, A. Timokhin, A. K. Harding and D. Kazanas, *Three-dimensional Kinetic Pulsar Magnetosphere Models: Connecting to Gamma-Ray Observations*, *ApJ* **857** (Apr., 2018) 44, [1710.03170].
- [88] R. M. Crocker et al., *Diffuse Galactic antimatter from faint thermonuclear supernovae in old stellar populations*, 1607.03495.
- [89] FERMI-LAT collaboration, J. Ballet, T. Burnett, S. Digel and B. Lott, *Fermi Large Area Telescope Fourth Source Catalog Data Release 2*, 2005.11208.
- [90] J. M. Yao, R. N. Manchester and N. Wang, *A New Electron-density Model for Estimation of Pulsar and FRB Distances*, *ApJ* **835** (Jan., 2017) 29, [1610.09448].
- [91] S. Weinberg, *Cosmology*. Oxford University Press, 2008.
- [92] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela et al., *Planck2013 results. xvi. cosmological parameters*, *Astronomy & Astrophysics* **571** (Oct, 2014) A16.
- [93] F. Nogueras-Lara, R. Schödel, A. T. Gallego-Calvente, E. Gallego-Cano, B. Shahzamanian, H. Dong et al., *Early formation and recent starburst activity in the nuclear disk of the Milky Way*, *Nature Astronomy* **4** (Dec., 2019) 377–381, [1910.06968].
- [94] R. G. Carlberg and K. A. Innanen, *Galactic Chaos and the Circular Velocity at the Sun*, *Astronom. J.* **94** (Sept., 1987) 666.
- [95] K. Kuijken and G. Gilmore, *The mass distribution in the galactic disc. I - A technique to determine the integral surface mass density of the disc near the sun.*, *MNRAS* **239** (Aug., 1989) 571–603.
- [96] R. Schönrich, J. Binney and W. Dehnen, *Local kinematics and the local standard of rest*, *MNRAS* **403** (Apr., 2010) 1829–1833, [0912.3693].
- [97] A. Spitkovsky, *Time-dependent Force-free Pulsar Magnetospheres: Axisymmetric and Oblique Rotators*, *ApJ* **648** (Sept., 2006) L51–L54, [astro-ph/0603147].

- [98] H. Haario, E. Saksman and J. Tamminen, *An adaptive metropolis algorithm*, *Bernoulli* **7** (2001) 223–242.
- [99] W. K. Hastings, *Monte carlo sampling methods using markov chains and their applications*, *Biometrika* **57** (1970) 97–109.
- [100] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari and D. B. Rubin, *Bayesian Data Analysis*. CRC Press, 3 ed., 2013.
- [101] W. Cash, *Parameter estimation in astronomy through application of the likelihood ratio.*, *ApJ* **228** (Mar., 1979) 939–947.
- [102] A. C. Cameron and P. K. Trivedi, *Regression Analysis of Count Data*,. Cambridge University Press, 1998.
- [103] S. Watanabe, *Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory*, *arXiv e-prints* (Apr., 2010) arXiv:1004.2316, [1004.2316].
- [104] S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, *Annals Math. Statist.* **9** (1938) 60–62.
- [105] D. Foreman-Mackey, *corner.py: Scatterplot matrices in python*, *The Journal of Open Source Software* **1** (jun, 2016) 24.
- [106] A. Gelman, *Two simple examples for understanding posterior p-values whose distributions are far from uniform*, *Electronic Journal of Statistics* **7** (2013) 2995–2602.
- [107] J. Bland-Hawthorn and O. Gerhard, *The Galaxy in Context: Structural, Kinematic, and Integrated Properties*, *Annual Review of Astronomy and Astrophysics* **54** (Sept., 2016) 529–596, [1602.07702].
- [108] M. Betancourt, *Cruising the simplex: Hamiltonian Monte Carlo and the Dirichlet distribution*, in *American Institute of Physics Conference Series* (P. Goyal, A. Giffin, K. H. Knuth and E. Vrscay, eds.), vol. 1443 of *American Institute of Physics Conference Series*, pp. 157–164, May, 2012. 1010.3436. DOI.
- [109] F. Calore, M. Di Mauro, F. Donato, J. W. T. Hessels and C. Weniger, *Radio Detection Prospects for a Bulge Population of Millisecond Pulsars as Suggested by Fermi-LAT Observations of the Inner Galaxy*, *ApJ* **827** (Aug., 2016) 143, [1512.06825].

- [110] L. J. Chang, S. Mishra-Sharma, M. Lisanti, M. Buschmann, N. L. Rodd and B. R. Safdi, *Characterizing the nature of the unresolved point sources in the galactic center: An assessment of systematic uncertainties*, *Physical Review D* **101** (Jan, 2020) .
- [111] F. Calore, I. Cholis and C. Weniger, *Background Model Systematics for the Fermi GeV Excess*, *JCAP* **1503** (2015) 038, [1409.0042].
- [112] J.-P. Macquart and N. Kanekar, *On Detecting Millisecond Pulsars at the Galactic Center*, *ApJ* **805** (June, 2015) 172, [1504.02492].
- [113] R. Eatough, T. J. W. Lazio, J. Casanellas, S. Chatterjee, J. M. Cordes, P. B. Demorest et al., *Observing Radio Pulsars in the Galactic Centre with the Square Kilometre Array*, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, p. 45, Apr., 2015. 1501.00281.
- [114] K. M. Rajwade, D. R. Lorimer and L. D. Anderson, *Detecting pulsars in the Galactic Centre*, *MNRAS* **471** (Oct., 2017) 730–739, [1611.06977].
- [115] S. Johnston, M. Kramer, D. R. Lorimer, A. G. Lyne, M. McLaughlin, B. Klein et al., *Discovery of two pulsars towards the Galactic Centre*, *MNRAS* **373** (Nov., 2006) L6–L10, [astro-ph/0606465].
- [116] J. S. Deneva, J. M. Cordes and T. J. W. Lazio, *Discovery of Three Pulsars from a Galactic Center Pulsar Population*, *ApJ* **702** (Sept., 2009) L177–L181, [0908.1331].
- [117] L. G. Spitler, K. J. Lee, R. P. Eatough, M. Kramer, R. Karuppusamy, C. G. Bassa et al., *Pulse Broadening Measurements from the Galactic Center Pulsar J1745-2900*, *ApJ* **780** (Jan., 2014) L3, [1309.4673].
- [118] G. C. Bower, A. Deller, P. Demorest, A. Brunthaler, R. Eatough, H. Falcke et al., *The Angular Broadening of the Galactic Center Pulsar SGR J1745-29: A New Constraint on the Scattering Medium*, *ApJ* **780** (Jan., 2014) L2, [1309.4672].
- [119] F. Camilo, M. Kerr, P. S. Ray, S. M. Ransom, J. Sarkissian, H. T. Cromartie et al., *Parkes Radio Searches of Fermi Gamma-Ray Sources and Millisecond Pulsar Discoveries*, *ApJ* **810** (Sept., 2015) 85, [1507.04451].
- [120] S. Sanpa-arsa, *Searching for new millisecond pulsars with the GBT in Fermi unassociated sources*. PhD thesis, University of Virginia, 7, 2016.
- [121] C. Ng, L. Guillemot, P. C. C. Freire, M. Kramer, D. J. Champion, I. Cognard et al., *A Shapiro delay detection in the pulsar binary system PSR J1811-2405*, *MNRAS* **493** (Mar., 2020) 1261–1267, [2002.01009].



- [122] J. R. Hurley, C. A. Tout, D. T. Wickramasinghe, L. Ferrario and P. D. Kiel, *Formation of binary millisecond pulsars by accretion-induced collapse of white dwarfs*, *Mon. Not. Roy. Astron. Soc.* **402** (02, 2010) 1437–1448.
- [123] A. J. Ruiter, L. Ferrario, K. Belczynski, I. R. Seitenzahl, R. M. Crocker and A. I. Karakas, *On the formation of neutron stars via accretion-induced collapse in binaries*, *Mon. Not. Roy. Astron. Soc.* **484** (Mar., 2019) 698–711, [1802.02437].
- [124] C. Fryer, W. Benz, M. Herant and S. A. Colgate, *What can the accretion-induced collapse of white dwarfs really explain?*, *Astrophys. J.* **516** (may, 1999) 892–899.
- [125] F. S. Kitaura, H.-T. Janka and W. Hillebrandt, *Explosions of o-ne-mg cores, the crab supernova, and subluminal type ii-p supernovae*, *Astronomy & Astrophysics* **450** (Apr, 2006) 345–350.
- [126] A. G. Lyne and D. R. Lorimer, *High birth velocities of radio pulsars*, *Nature* **369** (May, 1994) 127–129.
- [127] C. Eckner, X. Hou, P. D. Serpico, M. Winter, G. Zaharijas, P. Martin et al., *Millisecond Pulsar Origin of the Galactic Center Excess and Extended Gamma-Ray Emission from Andromeda: A Closer Look*, *ApJ* **862** (July, 2018) 79, [1711.05127].
- [128] J. Bédorf, E. Gaburov and S. Portegies Zwart, *A sparse octree gravitational N-body code that runs entirely on the GPU processor*, *Journal of Computational Physics* **231** (Apr., 2012) 2825–2839, [1106.1900].
- [129] M. S. Fujii, J. Bédorf, J. Baba and S. Portegies Zwart, *Modelling the Milky Way as a dry Galaxy*, *MNRAS* **482** (Jan., 2019) 1983–2015, [1807.10019].
- [130] K. Kuijken and J. Dubinski, *Nearly Self-Consistent Disc / Bulge / Halo Models for Galaxies*, *MNRAS* **277** (Dec., 1995) 1341.
- [131] L. M. Widrow and J. Dubinski, *Equilibrium Disk-Bulge-Halo Models for the Milky Way and Andromeda Galaxies*, *ApJ* **631** (Oct., 2005) 838–855, [astro-ph/0506177].
- [132] L. M. Widrow, B. Pym and J. Dubinski, *Dynamical Blueprints for Galaxies*, *ApJ* **679** (June, 2008) 1239–1259, [0801.3414].
- [133] F. Verbunt, A. Igoshev and E. Cator, *The observed velocity distribution of young pulsars*, *A&A* **608** (Dec., 2017) A57, [1708.08281].

- [134] L. Hernquist, *An Analytical Model for Spherical Galaxies and Bulges*, *ApJ* **356** (June, 1990) 359.
- [135] D. Foreman-Mackey, D. W. Hogg, D. Lang and J. Goodman, *emcee: The MCMC Hammer*, *Publications of the Astronomical Society of the Pacific* **125** (Mar., 2013) 306, [1202.3665].
- [136] C. ter Braak, *A markov chain monte carlo version of the genetic algorithm differential evolution: Easy bayesian computing for real parameter spaces*, *Statistics and Computing* **16** (09, 2006) 239–249.
- [137] C. ter Braak and J. Vrugt, *Differential evolution markov chain with snooker updater and fewer chains*, *Statistics and Computing* **18** (12, 2008) 435–446.
- [138] D. Syer and S. Tremaine, *Made-to-measure N-body systems*, *MNRAS* **282** (Sept., 1996) 223–233, [astro-ph/9605061].
- [139] F. de Lorenzi, V. P. Debattista, O. Gerhard and N. Sambhus, *NMAGIC: a fast parallel implementation of a  $\chi^2$ -made-to-measure algorithm for modelling observational data*, *MNRAS* **376** (Mar., 2007) 71–88, [astro-ph/0701582].
- [140] D. M. Nataf, A. Udalski, A. Gould, P. Fouqué and K. Z. Stanek, *The Split Red Clump of the Galactic Bulge from OGLE-III*, *Astrophys. J. lett.* **721** (Sept., 2010) L28–L32, [1007.5065].
- [141] A. McWilliam and M. Zoccali, *Two Red Clumps and the X-shaped Milky Way Bulge*, *ApJ* **724** (Dec., 2010) 1491–1502, [1008.0519].
- [142] M. Portail, C. Wegg, O. Gerhard and I. Martinez-Valpuesta, *Made-to-measure models of the Galactic box/peanut bulge: stellar and total mass in the bulge region*, *MNRAS* **448** (Mar., 2015) 713–731, [1502.00633].
- [143] D. Song, O. Macias and S. Horiuchi, *Inverse Compton emission from millisecond pulsars in the Galactic bulge*, *Phys. Rev. D* **99** (June, 2019) 123020, [1901.07025].
- [144] O. Macias, H. van Leijen, D. Song, S. Ando, S. Horiuchi and R. M. Crocker, *Cherenkov telescope array sensitivity to the putative millisecond pulsar population responsible for the galactic center excess*, *Monthly Notices of the Royal Astronomical Society* (May, 2021) .

# Appendix A

## Likelihood probability density function of resolved MSPs

In this appendix we implicitly assume that all probabilities are conditioned on the parameters ( $\theta$ ). Also, as all of our expression here are also for an individual MSP we leave the subscript on each observational quantity as implicit. As an example of these conventions,  $p(\text{obs}|l, b, F)$  is equivalent to  $p(\text{obs}|l_i, b_i, F_i, \theta)$ .

We assume the likelihood of observed MSPs with/without a parallax distance measurement depends only on distance  $d$ :

$$p(\text{obs}, l, b, d, P, \dot{P}, \mu_l, \mu_b, F, E_{\text{cut}}, \Gamma, \text{parallax/not parallax}) = p(\text{obs}, l, b, d, P, \dot{P}, \mu_l, \mu_b, F, E_{\text{cut}}, \Gamma) \times p(\text{parallax/not parallax}|d) \quad (\text{A.1})$$

where the probability of a parallax measurement given distance  $d$  is given by Eq. 2.47. In terms of the various components of the MSP model, the probability density function of resolved MSPs at  $l, b, d, P, \dot{P}, \mu_l, \mu_b, F, E_{\text{cut}}$  and  $\Gamma$  is:

$$p(\text{obs}, l, b, d, P, \dot{P}, \mu_l, \mu_b, F, E_{\text{cut}}, \Gamma) = \int \int p(\text{obs}, l, b, d, P, \dot{P}, \mu_l, \mu_b, F, E_{\text{cut}}, \Gamma, \alpha, P_I) d\alpha dP_I \quad (\text{A.2})$$

where we have integrated over the unknown magnetic axis angle and initial period. Then:

$$p(\text{obs}, l, b, d, P, \dot{P}, \mu_l, \mu_b, F, E_{\text{cut}}, \Gamma, \alpha, P_I) = p(\text{obs}|l, b, F) p(F|l, b, d, P, \dot{P}, \mu_l, \mu_b, E_{\text{cut}}, \alpha) \times p(E_{\text{cut}}, \Gamma|l, b, d, P, \dot{P}, \mu_l, \mu_b) p(P, \dot{P}|l, b, d, \mu_l, \mu_b, \alpha, P_I) p(\mu_l, \mu_b|l, b, d) \times p(l, b, d) p(\alpha) p(P_I) \quad (\text{A.3})$$

The probability of observing an MSP with flux  $F$  at  $l$  and  $b$  is given by Eq. 2.34, so:

$$p(\text{obs}|l, b, F) = p(F_{\text{th}} \leq F|l, b) \quad (\text{A.4})$$

The probability density function of an MSP having flux  $F$  conditional upon its other parameters is, given  $\log(F) = \log(\eta) + f(\dots)$  where  $f(\dots)$  is some function which doesn't depend on  $\eta$ :

$$\begin{aligned} p(F|l, b, d, P, \dot{P}, \mu_l, \mu_b, E_{\text{cut}}, \alpha) &= p(\eta|l, b, d, P, \dot{P}, \mu_l, \mu_b, E_{\text{cut}}, \alpha) \frac{\partial \eta}{\partial F} \\ &= p(\eta|l, b, d, P, \dot{P}, \mu_l, \mu_b, E_{\text{cut}}, \alpha) \frac{\eta}{F} \end{aligned} \quad (\text{A.5})$$

After a change of variables from period  $P$  and observed period derivative  $\dot{P}$  to magnetic field strength  $B$  and age  $t$ , we get:

$$p(P, \dot{P}|l, b, d, \mu_l, \mu_b, \alpha, P_l) = p(B, t|l, b, d, \alpha, P_l) \left| \frac{\partial B}{\partial P} \frac{\partial t}{\partial \dot{P}} - \frac{\partial B}{\partial \dot{P}} \frac{\partial t}{\partial P} \right| \quad (\text{A.6})$$

where to evaluate the Jacobian in the above equation we rewrite Eq. 2.21 as

$$\dot{P}_{\text{Shklovskii}} = C_1 P \quad (\text{A.7})$$

where  $C_1$  is a term that is independent of  $P$  and  $\dot{P}$ . We also rewrite Eq. 2.22 as

$$\dot{P}_{\text{Galactic}} = C_2 P \quad (\text{A.8})$$

where  $C_2$  is also independent of  $P$  and  $\dot{P}$ . We then obtain an equation for  $B$  in terms of  $P$  and  $\dot{P}$  using the above two equations with Eqs. 2.20 and 2.18 to get

$$B^2 = \frac{c^3 I P (-P(C_1 + C_2) + \dot{P})}{\pi^2 R^6 (\sin^2(\alpha) + 1)} \quad (\text{A.9})$$

Next we obtain an equation for  $t$  in terms of  $P$  and  $\dot{P}$  by substituting the above equation into Eq. 2.23 and solving  $t$  to get

$$t = \frac{P_l^2 - P^2}{2P((C_1 + C_2)P - \dot{P})} \quad (\text{A.10})$$

Using the above two equations we can then solve for the Jacobian term in Eq. A.6 to get

$$\left| \frac{\partial B}{\partial P} \frac{\partial t}{\partial \dot{P}} - \frac{\partial B}{\partial \dot{P}} \frac{\partial t}{\partial P} \right| = \frac{c^6 I^2 P^2}{2\pi^4 R^{12} (1 + \sin^2(\alpha))^2 B^3} \quad (\text{A.11})$$

where we have used Eq. A.9 to eliminate  $C_1$  and  $C_2$ . Similarly, using Eq. 2.11,  $\partial v / \partial \mu \propto d$ , so we find for proper motion:

$$p(\mu_l, \mu_b | l, b, d) \propto p(v_l, v_b | l, b, d) d^2 \quad (\text{A.12})$$

where the proportionality constant is independent of our parameters and data and so does not affect our results. For position:

$$p(l, b, d) \propto p(x, y, z) d^2 \cos(b) \quad (\text{A.13})$$

where  $d^2 \cos(b)$  is proportional to the Jacobian of the change of variables from  $x, y$  and  $z$  to  $l, b$  and  $d$ :

$$\begin{aligned} x &= -R_0 + d \cos(l) \cos(b) \\ y &= d \sin(l) \cos(b) \\ z &= d \sin(b) \end{aligned} \quad (\text{A.14})$$

The density of resolved MSPs is the sum of the density in the disk, boxy bulge and nuclear bulge populations:

$$\rho(\text{obs}, \dots) = N_{\text{disk}} p_{\text{disk}}(\text{obs}, \dots) + N_{\text{bb}} p_{\text{bb}}(\text{obs}, \dots) + N_{\text{nb}} p_{\text{nb}}(\text{obs}, \dots) \quad (\text{A.15})$$

where we can calculate the number in each population using the parameters  $\lambda_{\text{res}}$ ,  $\log_{10}(N_{\text{disk}}/N_{\text{bulge}})$  and  $\log_{10}(N_{\text{nb}}/N_{\text{bb}})$  and solving with:

$$\lambda_{\text{res}} = N_{\text{disk}} p_{\text{disk}}(\text{obs}) + N_{\text{bb}} p_{\text{bb}}(\text{obs}) + N_{\text{nb}} p_{\text{nb}}(\text{obs}) \quad (\text{A.16})$$

where each  $p(\text{obs})$  is evaluated using Eq. 2.43 for the corresponding spatial distribution.

In evaluating the likelihood for a given set of parameters, we used importance sampling to estimate integrals. As an example, in the case of the integrals over  $P_l$  and  $\alpha$  in Eq. A.2 this

integral becomes:

$$\begin{aligned}
 p(\text{obs}, l, b, d, P, \dot{P}, \mu_l, \mu_b, F, E_{\text{cut}}, \Gamma) &= \int \int p(\text{obs}, l, b, d, P, \dot{P}, \mu_l, \mu_b, F, E_{\text{cut}}, \Gamma, \alpha, P_I) d\alpha dP_I \\
 &= \int \int p(\text{obs}, l, b, d, P, \dot{P}, \mu_l, \mu_b, F, E_{\text{cut}}, \Gamma | \alpha, P_I) p(\alpha) p(P_I) d\alpha dP_I \\
 &\approx \frac{1}{N} \sum_{i=1}^N p(\text{obs}, l, b, d, P, \dot{P}, \mu_l, \mu_b, F, E_{\text{cut}}, \Gamma | \alpha_i, P_{Ii})
 \end{aligned} \tag{A.17}$$

where we sum over  $N$  samples  $\alpha_i$  and  $P_{Ii}$  from the probability distributions  $p(\alpha)$  and  $p(P_I)$ .

# Appendix B

## Measurement Uncertainties

For the measurements of  $F$ ,  $E_{\text{cut}}$  and  $\Gamma$  we use the 4FGL covariance matrices<sup>1</sup> for the uncertainty in  $N_0$ ,  $\gamma_1$  and  $a$ , with the spectrum of the form:<sup>2</sup>

$$\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^{\gamma_1} \exp(-aE^{\gamma_2}) \quad (\text{B.1})$$

where  $\gamma_2 = 2/3$ , and  $E_0$  is fixed at a different value for each MSP.

For the small number of MSPs in 4FGL-DR2 that were not fitted with spectra in the form of Eq. B.1, we simply use the reported flux and its error, treating  $E_{\text{cut}}$  and  $\Gamma$  as missing. Since the catalog required  $\Gamma > 0$ , we also treat  $E_{\text{cut}}$  and  $\Gamma$  as missing in the case where some MSPs had  $\Gamma$  fixed near 0. If, based on central estimates of  $\dot{P}$ ,  $\mu$  and  $d$ ,  $\dot{P}_{\text{int}}$  appears to be negative for a particular MSP (i.e., the MSP is apparently spinning up) we treat  $\dot{P}$  as missing. An apparently negative intrinsic period derivatives can result in the case that the true proper motion and/or distance is lower than the value we use, therefore causing the Shklovskii effect to be overestimated. An example of an updated proper motion measurement fixing this issue is given for PSR J1231-1411 in Abdo et al. [45]. A second potential cause of negative period derivative is radial acceleration in the direction of the sun significantly in excess of that accounted for by the  $\dot{P}_{\text{Galactic}}$  term in equations 2.20 and 2.22.

In the case of distance uncertainties, for the dispersion measure we use a method similar to that of Bartels et al. [16]: The relative uncertainty in the dispersion measure is typically very small, so we assume that it is measured exactly. The main source of distance uncertainty is, therefore, associated with the model of the free electron density along the line of sight for each MSP. We use the YMW16 model of Yao et al. [90] and integrate over the uncertainty in

---

<sup>1</sup>Kindly provided to us by Dr Jean Ballet of the Fermi-LAT collaboration.

<sup>2</sup>"PLSuperExpCutoff2" at [https://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/source\\_models.html](https://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/source_models.html)

the model parameters  $\theta_{\text{YMW16}}$  so that:

$$\begin{aligned}
 p(\dots, \text{DM}_j, \dots) &= \int d\theta_{\text{YMW16}} p(\theta_{\text{YMW16}}) p(\dots, \text{DM}_j, \dots | \theta_{\text{YMW16}}) \\
 &= \int d\theta_{\text{YMW16}} p(\theta_{\text{YMW16}}) \\
 &\quad \times \left| \frac{\partial}{\partial \text{DM}} d(\text{DM}_j, \theta_{\text{YMW16}}) \right| p(\dots, d(\text{DM}_j, \theta_{\text{YMW16}}), \dots) \\
 &= \int d\theta_{\text{YMW16}} \frac{p(\theta_{\text{YMW16}})}{n_e(d(\text{DM}_j, \theta_{\text{YMW16}}), \theta_{\text{YMW16}})} p(\dots, d(\text{DM}_j, \theta_{\text{YMW16}}), \dots)
 \end{aligned} \tag{B.2}$$

where  $d(\text{DM}, \theta_{\text{YMW16}})$  is distance as a function of dispersion measure and YMW16 model parameters. This can be found by solving for  $d$  in  $\text{DM}(\theta_{\text{YMW16}}) = \int_0^d n_e(s, \theta_{\text{YMW16}}) ds$  where  $n_e(s, \theta_{\text{YMW16}})$  is the free electron density at distance  $s$  for model parameters  $\theta_{\text{YMW16}}$ .

From Eqs. 2.45, 2.48 and A.15 we can work out the fraction of the likelihood contributed by the boxy bulge and nuclear bulge components for resolved MSP  $i$  given the model parameters and data uncertainties:

$$p(\text{Bulge MSP}_i) = \frac{\rho_{\text{bb}}(\mathbf{D}_i) + \rho_{\text{nb}}(\mathbf{D}_i)}{\rho(\mathbf{D}_i)} \tag{B.3}$$

where  $\rho_{\text{bb}}(\mathbf{D}_i) = N_{\text{bb}} p_{\text{bb}}(\text{obs}, \mathbf{D}_i)$  and  $\rho_{\text{nb}}(\mathbf{D}_i) = N_{\text{nb}} p_{\text{nb}}(\text{obs}, \mathbf{D}_i)$ .



# Appendix C

## Sampling methods

The disk model, given in Eq. 2.1, can be sampled from by sampling in cylindrical coordinates a random  $R$ ,  $z$  and  $\phi$ . The radial coordinate  $R$  is drawn from:

$$p(R) = \frac{1}{\sigma_r^2} \exp(-R^2/2\sigma_r^2) R \quad (\text{C.1})$$

for which, using the inverse of the cumulative distribution function of  $R$ , if  $u$  is a uniformly random draw from  $[0, 1]$ :

$$R = \sigma_r \sqrt{-2 \log(1 - u)} \quad (\text{C.2})$$

The height  $z$  is drawn from:

$$p(z) = \frac{1}{2z_0} \exp(-|z|/z_0) \quad (\text{C.3})$$

which can be done by drawing  $|z|$  from an exponential distribution and choosing either a positive or negative sign each with 0.5 probability. Finally,  $\phi$  is drawn from a uniform distribution on  $[0, 2\pi]$ . For the boxy bulge and nuclear bulge distributions, we sampled the density using MCMC.

To sample from the age distribution, we used MCMC. We used standard library functions to sample from the various Gaussian distributions. To sample  $\alpha$ , again using inverse transform sampling:

$$\alpha = \arccos(1 - 2u) \quad (\text{C.4})$$

for  $u$  uniformly drawn from  $[0, 1]$ .



## Appendix D

# Watanabe-Akaike Information Criterion (WAIC)

In order to rank the various models of the Galactic MSP population in Chapter 2, correcting for the varying number of parameters, we use the WAIC. In this appendix we define the WAIC, then derive the contribution to the WAIC associated with the resolved MSPs component of the likelihood.

The WAIC is defined in terms of the log pointwise predictive density (lppd) and effective number of parameters ( $p_{\text{WAIC}}$ ) as [100]:

$$\text{WAIC} = -2(\text{lppd} - p_{\text{WAIC}}) \quad (\text{D.1})$$

where for data  $\mathbf{y}_1, \dots, \mathbf{y}_N$  and  $S$  parameter sets  $\boldsymbol{\theta}^s$  in our Markov chain:

$$\text{lppd} = \sum_{i=1}^N \log \left( \frac{1}{S} \sum_{s=1}^S p(\mathbf{y}_i | \boldsymbol{\theta}^s) \right) \quad (\text{D.2})$$

where  $p(\mathbf{y}_i | \boldsymbol{\theta}^s)$  is the predictive density of  $\mathbf{y}_i$  given model parameters  $\boldsymbol{\theta}^s$ . Using the WAIC1 option from Gelman et al. [100]

$$p_{\text{WAIC}} = 2 \sum_{i=1}^N \left( \log \left( \frac{1}{S} \sum_{s=1}^S p(\mathbf{y}_i | \boldsymbol{\theta}^s) \right) - \frac{1}{S} \sum_{s=1}^S \log(p(\mathbf{y}_i | \boldsymbol{\theta}^s)) \right) \quad (\text{D.3})$$

We can write the WAIC as the sum of two components,  $\text{WAIC}_{\text{res}}$  and  $\text{WAIC}_{\text{GCE}}$ . For the GCE contribution,  $\text{WAIC}_{\text{GCE}}$ , we use the Gaussian likelihood for each bin as in Eq. 2.50. If we use, for  $\text{WAIC}_{\text{res}}$ , the contribution of resolved MSPs, a Poisson distribution for bins in the

several dimensions in which we have data, then define  $\lambda_{i,s}$  as the expectation value for bin  $i$  for parameter set  $s$  and  $n_i$  the number of observations in bin  $i$ , then:

$$\begin{aligned} p(n_i | \lambda_{i,s}) &= \frac{\exp(-\lambda_{i,s}) \lambda_{i,s}^{n_i}}{n_i!} \\ &= \frac{\exp(-\delta \rho_{i,s}) (\delta \rho_{i,s})^{n_i}}{n_i!} \end{aligned} \quad (\text{D.4})$$

where  $\lambda_{i,s} = \delta \rho_{i,s}$  with  $\delta$  the bin volume and  $\rho_{i,s}$  the average density within bin  $i$ . Then if we choose  $\delta$  small enough such that  $\delta \rho_{i,s} \ll 1$  and  $n_i$  is either 0 or 1, we can derive the contribution to the lppd from the resolved MSP data:

$$\begin{aligned} \text{lppd}_{\text{res}} &= \sum_{i=1}^N \log \left( \frac{1}{S} \sum_{s=1}^S p(n_i | \lambda_{i,s}) \right), \text{ using Eq. D.2} \\ &= -N \log(S) \\ &\quad + \sum_{i=1}^N \log \left( \sum_{s=1}^S \exp(-\delta \rho_{i,s}) (\delta \rho_{i,s})^{n_i} \right), \text{ using Eq. D.4 and that } n_i \text{ is either 0 or 1} \\ &\approx -N \log(S) + \sum_{i=1}^N \log \left( \sum_{s=1}^S (1 - \delta \rho_{i,s}) (\delta \rho_{i,s})^{n_i} \right), \text{ using } \delta \rho_{i,s} \ll 1 \\ &= -N \log(S) + \sum_{i=1}^N \begin{cases} \log(\sum_{s=1}^S (1 - \delta \rho_{i,s})) & n_i = 0 \\ \log(\sum_{s=1}^S (1 - \delta \rho_{i,s}) (\delta \rho_{i,s})) & n_i = 1 \end{cases} \\ &\approx -N \log(S) + \sum_{i=1}^N \begin{cases} \log(S(1 - \frac{1}{S} \sum_{s=1}^S \delta \rho_{i,s})) & n_i = 0 \\ \log(\sum_{s=1}^S \delta \rho_{i,s}) & n_i = 1, \text{ using } \delta \rho_{i,s} \ll 1 \end{cases} \\ &= -N \log(S) + \sum_{i=1}^N \begin{cases} (\log(S) + \log(1 - \frac{1}{S} \sum_{s=1}^S \delta \rho_{i,s})) & n_i = 0 \\ (\log(\delta) + \log(\sum_{s=1}^S \rho_{i,s})) & n_i = 1 \end{cases} \\ &= -N \log(S) + N_{n_i=0} \log(S) + N_{n_i=1} \log(\delta) + \sum_{i=1}^N \begin{cases} \log(1 - \frac{1}{S} \sum_{s=1}^S \delta \rho_{i,s}) & n_i = 0 \\ \log(\sum_{s=1}^S \rho_{i,s}) & n_i = 1 \end{cases} \end{aligned} \quad (\text{D.5})$$

where  $N_{n_i=0}$  is the number of bin with zero counts and  $N_{n_i=1}$  for the number of bins with one count.

Using  $-N + N_{n_i=0} = -N_{n_i=1}$  in Eq. D.5 gives

$$\begin{aligned}
 \text{lppd}_{\text{res}} &= -N_{n_i=1} \log(S) + N_{n_i=1} \log(\delta) + \sum_{i=1}^N \begin{cases} \log(1 - \frac{1}{S} \sum_{s=1}^S \delta \rho_{i,s}) & n_i = 0 \\ \log(S \frac{1}{S} \sum_{s=1}^S \rho_{i,s}) & n_i = 1 \end{cases} \\
 &\approx -N_{n_i=1} \log(S) \\
 &\quad + N_{n_i=1} \log(\delta) + \sum_{i=1}^N \begin{cases} -\frac{1}{S} \sum_{s=1}^S \delta \rho_{i,s} & n_i = 0, \text{ using } \delta \rho_{i,s} \ll 1 \\ (\log(S) + \log(\frac{1}{S} \sum_{s=1}^S \rho_{i,s})) & n_i = 1 \end{cases} \\
 &= N_{n_i=1} \log(\delta) + \sum_{i=1}^N \begin{cases} -\frac{1}{S} \sum_{s=1}^S \delta \rho_{i,s} & n_i = 0 \\ \log(\frac{1}{S} \sum_{s=1}^S \rho_{i,s}) & n_i = 1 \end{cases}
 \end{aligned} \tag{D.6}$$

As most voxels will have  $n_i = 0$  we can approximate  $\lambda_s = \sum_{n_i=0, n_i=1} \delta \rho_{i,s} \approx \sum_{n_i=0} \delta \rho_{i,s}$  where  $\lambda_s$  is simply the total expected number of resolved MSPs parameter. Therefore,

$$\begin{aligned}
 \text{lppd}_{\text{res}} &\approx N_{n_i=1} \log(\delta) - \frac{1}{S} \sum_{s=1}^S \lambda_s + \sum_{i=1}^N \begin{cases} 0 & n_i = 0 \\ \log(\frac{1}{S} \sum_{s=1}^S \rho_{i,s}) & n_i = 1 \end{cases} \\
 &= N_{n_i=1} \log(\delta) - \frac{1}{S} \sum_{s=1}^S \lambda_s + \sum_{j=1}^{N_{\text{res}}} \log \left( \frac{1}{S} \sum_{s=1}^S N_{\text{tot},s} p(\text{obs}, \mathbf{D}_j \mid \boldsymbol{\theta}^s) \right)
 \end{aligned} \tag{D.7}$$

where in the last line we have used Eq. 2.45. The first term in the above equation can be ignored as long as we are comparing models fitted using the same data.

The other term we need to evaluate is given by Eq. D.3 which we write in our notation as

$$p_{\text{WAIC}, \text{res}} = 2 \sum_{i=1}^N \left( \log \left( \frac{1}{S} \sum_{s=1}^S p(n_i \mid \lambda_{i,s}) \right) - \frac{1}{S} \sum_{s=1}^S \log(p(n_i \mid \lambda_{i,s})) \right). \tag{D.8}$$

We substitute Eq. D.4 to get

$$\begin{aligned}
 p_{\text{WAIC}, \text{res}} &= -2N \log(S) \\
 &\quad + 2 \sum_{i=1}^N \left( \log \left( \sum_{s=1}^S \exp(-\delta \rho_{i,s}) (\delta \rho_{i,s})^{n_i} \right) - \frac{1}{S} \sum_{s=1}^S \log(\exp(-\delta \rho_{i,s}) (\delta \rho_{i,s})^{n_i}) \right)
 \end{aligned} \tag{D.9}$$

where we have used the fact that  $n_i = 0$  or  $n_i = 1$  so  $n_i! = 1$  in either case. Next we separate out the two possible values for  $n_i$  and make use of  $\delta\rho_{i,s} \ll 1$  as follows:

$$\begin{aligned}
p_{\text{WAIC, res}} &= -2N \log(S) \\
&\quad + 2 \sum_{i=1}^N \begin{cases} (\log(\sum_{s=1}^S \exp(-\delta\rho_{i,s})) - \frac{1}{S} \sum_{s=1}^S \log(\exp(-\delta\rho_{i,s}))) & n_i = 0 \\ (\log(\sum_{s=1}^S \exp(-\delta\rho_{i,s})(\delta\rho_{i,s})) - \frac{1}{S} \sum_{s=1}^S \log(\exp(-\delta\rho_{i,s})(\delta\rho_{i,s}))) & n_i = 1 \end{cases} \\
&\approx -2N \log(S) \\
&\quad + 2 \sum_{i=1}^N \begin{cases} (\log(\sum_{s=1}^S (1 - \delta\rho_{i,s})) - \frac{1}{S} \sum_{s=1}^S (-\delta\rho_{i,s})) & n_i = 0 \\ (\log(\sum_{s=1}^S (1 - \delta\rho_{i,s})(\delta\rho_{i,s})) - \frac{1}{S} \sum_{s=1}^S \log((1 - \delta\rho_{i,s})(\delta\rho_{i,s}))) & n_i = 1 \end{cases} \\
&\approx -2N \log(S) + 2 \sum_{i=1}^N \begin{cases} (\log(S(1 - \frac{1}{S} \sum_{s=1}^S \delta\rho_{i,s})) - \frac{1}{S} \sum_{s=1}^S (-\delta\rho_{i,s})) & n_i = 0 \\ (\log(S \frac{1}{S} \sum_{s=1}^S \delta\rho_{i,s}) - \frac{1}{S} \sum_{s=1}^S \log(\delta\rho_{i,s})) & n_i = 1 \end{cases} \\
&= -2N \log(S) + 2 \sum_{i=1}^N \begin{cases} (\log(S) + \log(1 - \frac{1}{S} \sum_{s=1}^S \delta\rho_{i,s}) + \frac{1}{S} \sum_{s=1}^S \delta\rho_{i,s}) & n_i = 0 \\ (\log(S) + \log(\frac{1}{S} \sum_{s=1}^S \delta\rho_{i,s}) - \frac{1}{S} \sum_{s=1}^S \log(\delta\rho_{i,s})) & n_i = 1 \end{cases} \\
&\approx 2 \sum_{i=1}^N \begin{cases} (-\frac{1}{S} \sum_{s=1}^S \delta\rho_{i,s} + \frac{1}{S} \sum_{s=1}^S \delta\rho_{i,s}) & n_i = 0 \\ (\log(\frac{1}{S} \sum_{s=1}^S \delta\rho_{i,s}) - \frac{1}{S} \sum_{s=1}^S \log(\delta\rho_{i,s})) & n_i = 1 \end{cases} \\
&= 2 \sum_{i=1}^N \begin{cases} 0 & n_i = 0 \\ (\log(\frac{1}{S} \sum_{s=1}^S \rho_{i,s}) - \frac{1}{S} \sum_{s=1}^S \log(\rho_{i,s})) & n_i = 1. \end{cases}
\end{aligned} \tag{D.10}$$

Substituting Eq. 2.45 into the above equation gives

$$p_{\text{WAIC, res}} = 2 \sum_{j=1}^{N_{\text{res}}} \left( \log \left( \frac{1}{S} \sum_{s=1}^S N_{\text{tot},s} p(\text{obs}, \mathbf{D}_j \mid \boldsymbol{\theta}^s) \right) - \frac{1}{S} \sum_{s=1}^S \log(N_{\text{tot},s} p(\text{obs}, \mathbf{D}_j \mid \boldsymbol{\theta}^s)) \right). \tag{D.11}$$

Substituting Eqs. D.11 and D.7 into Eq. D.1 gives

$$\begin{aligned}
\text{WAIC}_{\text{res}} &= 2 \left[ \frac{1}{S} \sum_{s=1}^S \lambda_s - \sum_{j=1}^{N_{\text{res}}} \log \left( \frac{1}{S} \sum_{s=1}^S N_{\text{tot},s} p(\text{obs}, \mathbf{D}_j \mid \boldsymbol{\theta}^s) \right) \right] \\
&\quad + 4 \left\{ \sum_{j=1}^{N_{\text{res}}} \left[ \log \left( \frac{1}{S} \sum_{s=1}^S N_{\text{tot},s} p(\text{obs}, \mathbf{D}_j \mid \boldsymbol{\theta}^s) \right) - \frac{1}{S} \sum_{s=1}^S \log(N_{\text{tot},s} p(\text{obs}, \mathbf{D}_j \mid \boldsymbol{\theta}^s)) \right] \right\}.
\end{aligned} \tag{D.12}$$

# **Appendix E**

## **MWb and MWc0.8 *N*-body Simulation Figures**

As the results of the MWb and MWc0.8 *N*-body simulations in Chapter 3 are similar to the MWa case, we have moved many figures specific to those models to this appendix.

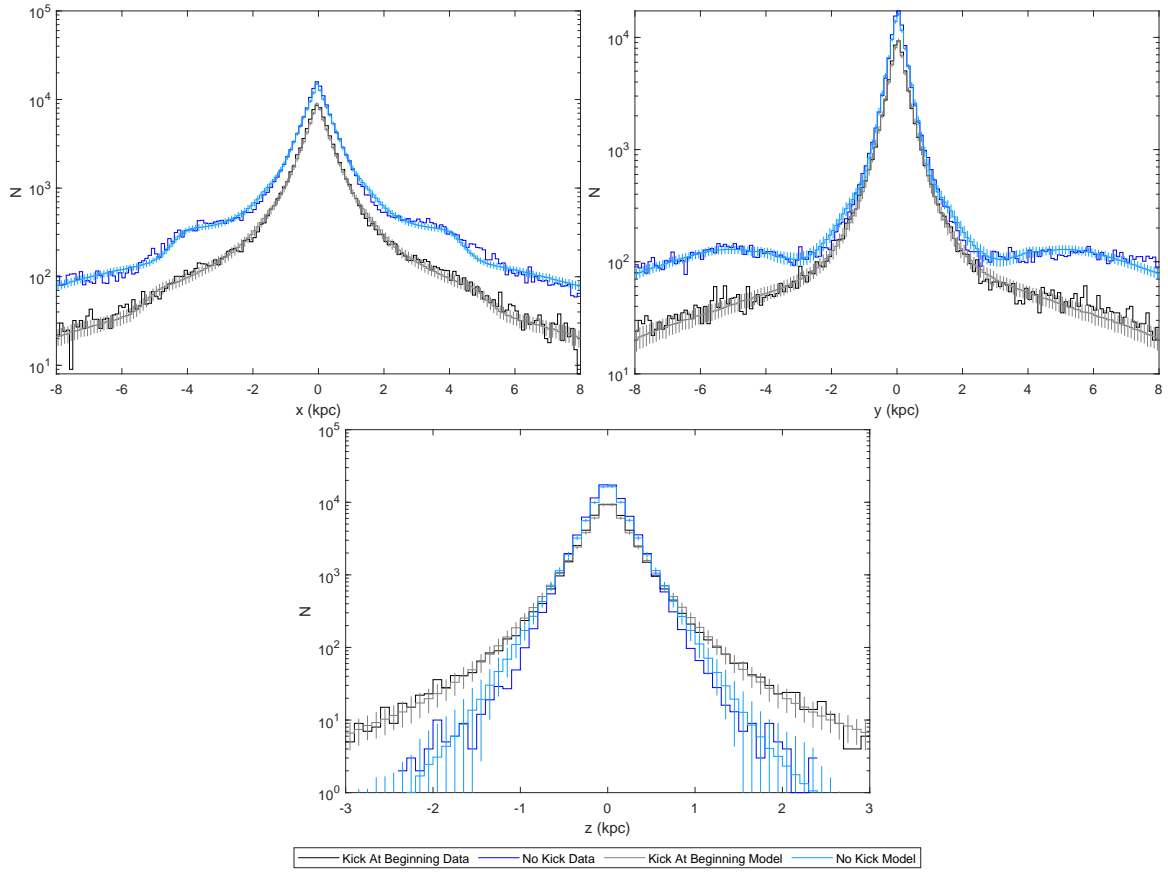


Fig. E.1 MWb profile along  $x$ ,  $y$  and  $z$  axes with kicks occurring at the beginning. We show both  $N$ -body simulation data and data simulated using the fitted model. For the fitted model we show the mean number of particles in each bin and the standard deviation.



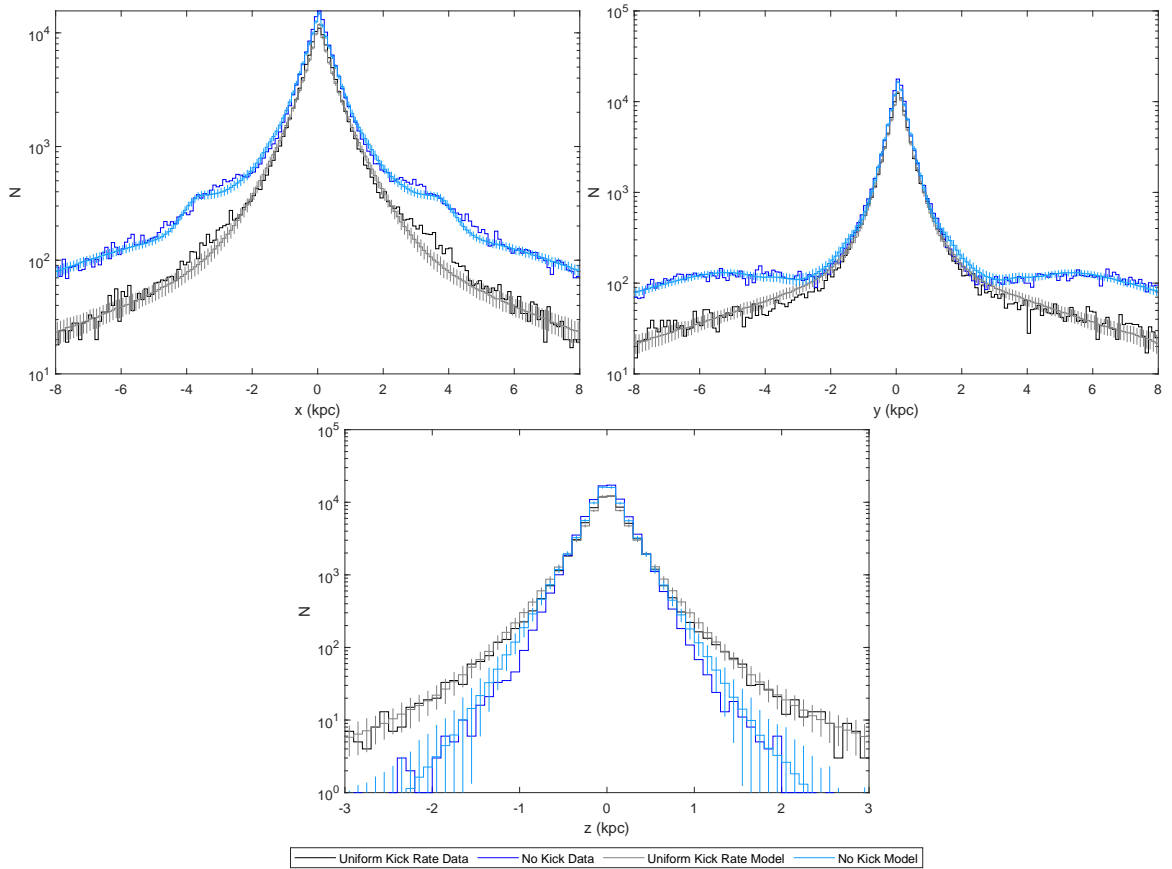


Fig. E.2 MWb profile along  $x$ ,  $y$  and  $z$  axes with a uniform kick rate. We show both  $N$ -body simulation data and data simulated using the fitted model. For the fitted model we show the mean number of particles in each bin and the standard deviation.

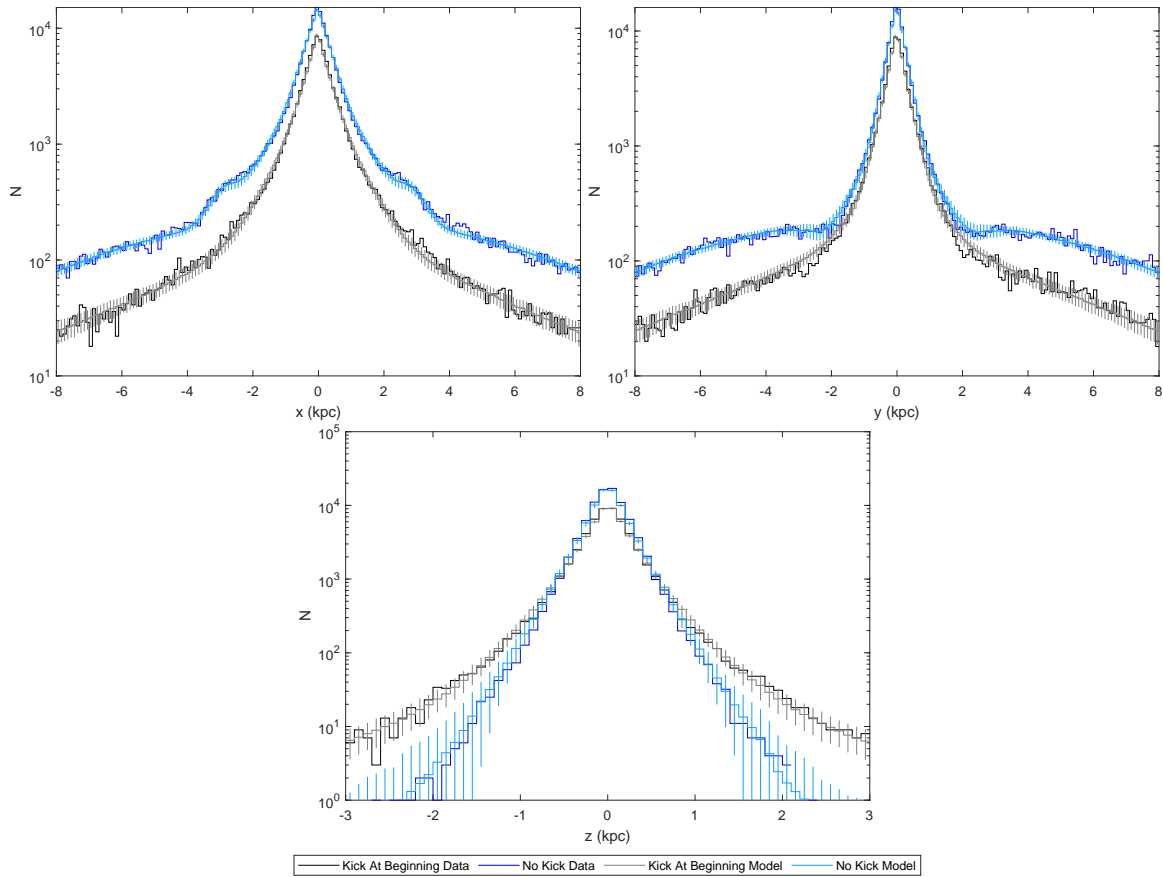


Fig. E.3 MWc0.8 profile along  $x$ ,  $y$  and  $z$  axes with kicks occurring at the beginning. We show both  $N$ -body simulation data and data simulated using the fitted model. For the fitted model we show the mean number of particles in each bin and the standard deviation.

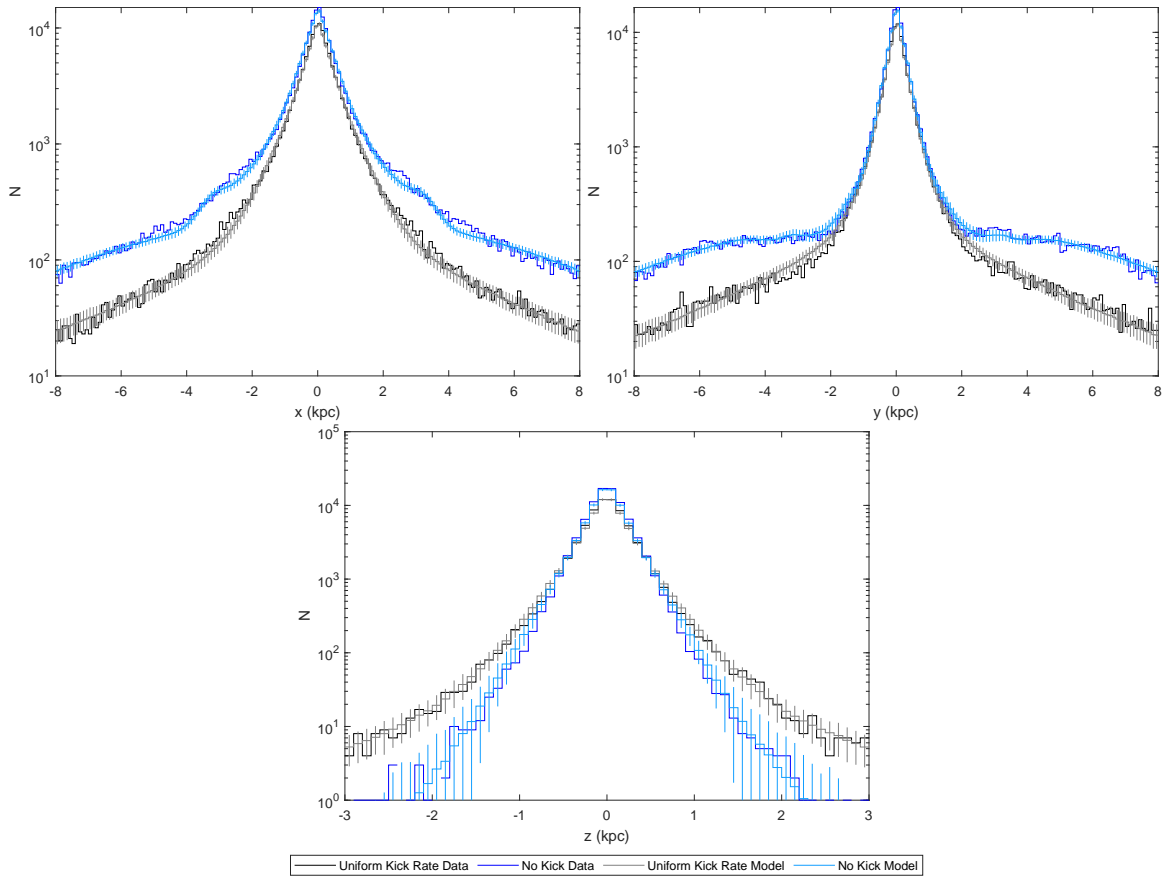


Fig. E.4 MWc0.8 profile along  $x$ ,  $y$  and  $z$  axes with a uniform kick rate. We show both  $N$ -body simulation data and data simulated using the fitted model. For the fitted model we show the mean number of particles in each bin and the standard deviation.

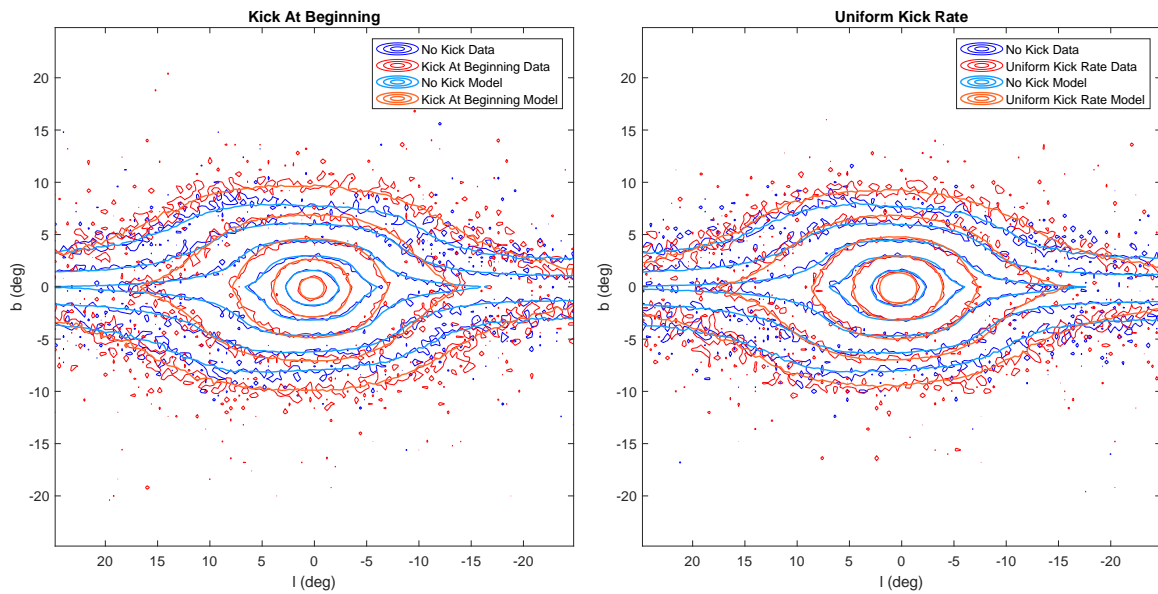


Fig. E.5 MWb flux distribution in Galactic coordinates. The contours for each distribution are at 1, 2, 4, 8 and 16 times the mean in this region. The Sun is placed at a distance of 7.9 kpc, at an angle relative to the bar of  $20^\circ$  and at a height of 15 pc.

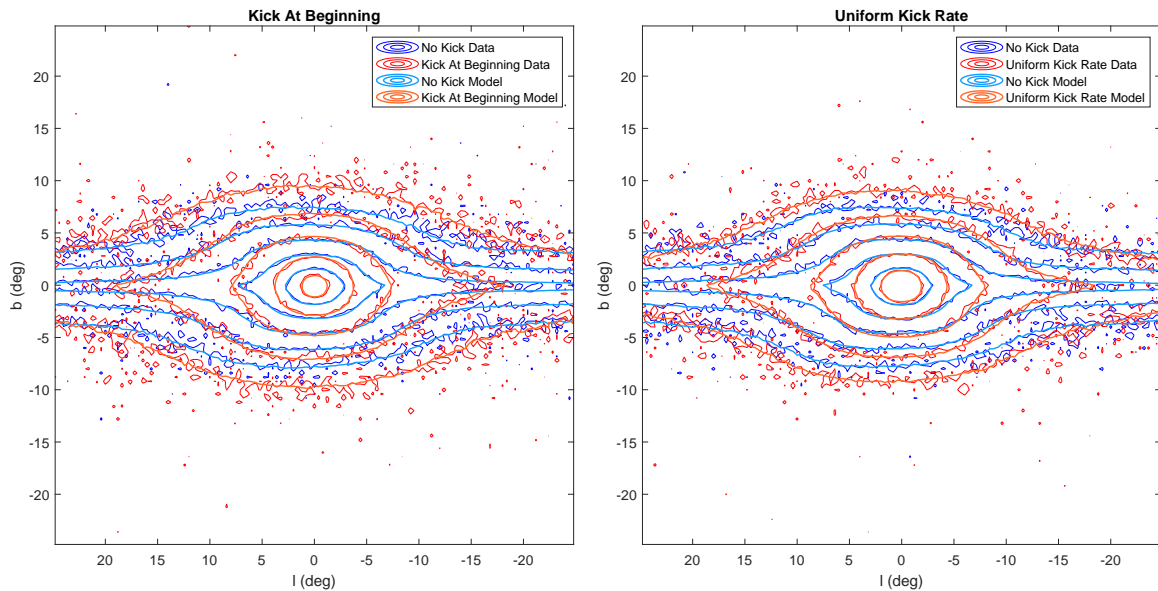


Fig. E.6 MWc0.8 flux distribution in Galactic coordinates. The contours for each distribution are at 1, 2, 4, 8 and 16 times the mean in this region. The Sun is placed at a distance of 7.9 kpc, at an angle relative to the bar of  $20^\circ$  and at a height of 15 pc.